AI and Social Aspects

Weekly Analysis — https://ainews.social

In Rotterdam, an algorithm flagged thousands of families for welfare fraud investigation based on their postal codes, languages spoken, and debts. The system, designed to protect public resources, instead created what journalists called a "suspicion machine" that systematically targeted immigrant neighborhoods and single mothers. This story, meticulously documented in [11], reveals a fundamental paradox in contemporary AI discourse: while we produce ever more analysis of algorithmic systems, the power to shape these technologies remains concentrated in the hands of those who already hold authority, while those most affected by algorithmic decisions remain structurally silenced.

[11] Inside the Suspicion Machine - WIRED

The discourse around AI and its social impacts has reached a fever pitch. Academic journals overflow with studies of bias, conferences dedicate entire tracks to AI ethics, and policy documents proliferate across institutions. Yet this explosion of discourse masks a troubling reality: the conversation itself has become a site of power reproduction. Those who speak loudest about AI—researchers at well-funded institutions, technology companies, policy makers in the Global North—are rarely those whose lives are reshaped by algorithmic decisions. The 39.5% of AI discourse focused on ethical failures might initially seem like progress, a sign that we're taking harm seriously. But this fixation on documenting problems without building solutions reveals something darker: a discourse that performs concern while preserving existing power structures.

This essay interrogates the power relations embedded in AI discourse itself. By examining who speaks, who is spoken about, and who remains silent, we can begin to understand how the very act of discussing AI's social impacts can reinforce the inequalities these discussions purport to address. The evidence reveals a discourse dominated by institutional voices from the Global North, technical optimization frameworks that sideline human experience, and a persistent gap between those who document harm and those who experience it. Most troublingly, while we excel at diagnosing algorithmic bias and documenting its impacts on marginalized communities, we consistently fail to translate this knowledge into meaningful change. This is not an accident but a feature of how power operates through discourse—

maintaining its authority by controlling both the conversation and its limits.

The Architecture of Algorithmic Authority

The question of who shapes AI systems and their deployment reveals a stark concentration of power in specific institutional locations. Major technology companies, government agencies, and academic institutions in the Global North dominate both the development of AI systems and the discourse about their impacts. This concentration is not merely about resources or technical capability—it represents a fundamental ordering of whose perspectives matter in determining how algorithmic systems intersect with human lives.

Consider the seemingly neutral domain of workplace AI adoption. When [7] reports on differential AI usage creating a "two-tier workforce," the framing itself reveals power dynamics. The narrative centers on productivity metrics and skill acquisition, treating the emerging inequality as a natural market outcome rather than a consequence of deliberate design choices. The workers who find themselves on the wrong side of this divide—those without access to training, time for experimentation, or jobs that allow for AI integration—become statistics in someone else's productivity study. Meanwhile, the companies deploying these systems shape both the technology and the narrative about its inevitability.

[7] Anthropic Data Shows AI Skills Gap Splitting Workforces

The educational sector provides another window into these power dynamics. When judges report broad AI adoption without corresponding training, as documented in [Federal judges report broad adoption of AI tools], we see authority figures empowered to use systems they don't fully understand to make decisions affecting human lives. The 45.5% of judges reporting no AI training represents not just a skills gap but a power gap—between those who design systems and those who must use them, between those who understand the technology's limitations and those who must trust its outputs. This dynamic repeats across sectors: educators required to use AI detection tools shown to be biased, welfare administrators implementing algorithms they cannot interrogate, healthcare workers relying on systems trained on data that excludes their patients' populations.

The concentration of AI development in specific geographic and institutional locations creates what researchers have termed "digital colonialism." As [16] argues, this represents a new form of economic and cultural domination where data is extracted from the Global South to train systems designed in the Global North, which are then

[16] Ramificações Do Colonialismo Digital: Inteligência Artificial Como a ...

deployed back onto Southern populations without regard for local contexts or needs. The African Union's continental AI strategy explicitly recognizes this dynamic, noting how AI systems developed elsewhere fail to account for African languages, educational contexts, and cultural values. When [19] documents how AI tutoring systems fail to recognize African knowledge systems or pedagogical approaches, we see power operating through technological design—determining whose ways of knowing count as knowledge worth encoding.

[19] The cultural cost of AI in Africa's education systems - UNESCO

The corporate capture of AI development extends beyond mere market dominance. [Big Tech sets unfair terms and conditions for AI data workers globally] reveals how the same companies positioning themselves as AI innovators create exploitative conditions for the workers who make their systems possible. Data labelers in Kenya, Venezuela, and the Philippines work for wages below local minimums, without job security or recourse to challenge unfair treatment. Their labor—categorizing images, flagging toxic content, teaching machines to recognize human faces—remains invisible in the glossy presentations about AI's transformative potential. This invisibility is not incidental but structural: acknowledging the human labor that enables AI would complicate narratives about autonomous, intelligent machines and raise uncomfortable questions about who benefits from AI development.

Even in public sector applications, where one might expect democratic oversight, power concentrates in technical teams and vendor relationships. The French CAF (Caisse d'Allocations Familiales) algorithm that [17] exposed through legal action demonstrates this dynamic. The algorithm, developed to detect welfare fraud, operated for years without meaningful oversight, scoring citizens based on opaque criteria. When advocates finally forced disclosure of the source code, they found a system that penalized poverty itself—flagging single parents, immigrants, and those with unstable housing as probable fraudsters. The power to define suspicion, to encode assumptions about who deserves public benefits, rested with algorithm designers whose biases became mathematical law.

[17] Scoring of welfare beneficiaries: the indecency of CAF's algorithm now ...

When Numbers Speak Louder Than Lives

The dominance of quantification in AI discourse reveals another dimension of power: the ability to transform human suffering into abstract metrics that obscure more than they reveal. When [AI and Welfare: A Failed Match] reports that algorithmic systems denied food rations to 47.3 million people in India, the sheer scale of the number

can paradoxically diminish its impact. These are not simply statistics but families who went hungry, children whose education suffered from malnutrition, elderly people who died waiting for benefits wrongly denied. Yet in the discourse of AI ethics, such impacts often become data points in discussions of "algorithmic bias" or "implementation challenges," language that maintains clinical distance from human consequences.

This transformation of lived experience into quantified harm serves specific power interests. For system designers and institutional deployers, numbers provide a shield against moral reckoning. A facial recognition system that misidentifies Black women at rates 35% higher than white men becomes a technical problem requiring algorithmic adjustment rather than a fundamental question about surveillance, racial justice, and the right to move through public space without suspicion. When [15], the individual story breaks through the statistical abstraction—a woman detained, her life disrupted, her faith in justice systems shattered. Yet even such vivid cases often get absorbed back into aggregate discussions of "error rates" and "confidence thresholds."

[15] Police used AI facial recognition to arrest a Tennessee woman for crimes committed in a state she says she's never visited

The quantification of educational inequality through AI metrics demonstrates this dynamic with particular clarity. [5] meticulously documents how AI systems in education perpetuate existing biases, showing differential impacts on students by race, class, and language. But the very act of studying bias through quantitative frameworks can obscure the qualitative experiences of students subjected to these systems. A child flagged as "at-risk" by an algorithm carries that label through their educational journey, shaping teacher expectations and administrative decisions in ways that no retrospective bias study can fully capture. The power to be seen as an individual rather than a data point—to have one's story heard rather than one's statistics analyzed—remains unequally distributed.

[5] Algorithmic Bias in Education | International Journal of Artificial ...

The workplace provides another arena where quantification obscures power relations. When [9] documents how algorithmic wage-setting systems have contributed to $8.7 billion in lost income for gig workers globally, the aggregate figure tells only part of the story. Behind each dollar lost is a driver who couldn't make rent, a delivery worker choosing between gas and groceries, a family slipping from precarity into poverty. The companies deploying these systems frame them as neutral "optimization" tools, using the language of efficiency to mask decisions about how to distribute value between shareholders and workers. The algorithm becomes a convenient scapegoat—"the system" determined your pay—obscuring the human choices about what to optimize and whose interests to prioritize.

[9] How companies are using AI to pay workers as little as possible

Even well-intentioned efforts to document algorithmic harm can participate in this numbers game. Research on AI bias often focuses on measurable disparities—differences in error rates, accuracy gaps between demographic groups, statistical evidence of discrimination. While such work is necessary, it can inadvertently accept the premise that fairness is primarily a mathematical property rather than a lived experience. As [2] argues, AI systems trained on biased historical data inevitably reproduce past inequalities. But focusing on data quality and algorithmic fairness metrics can obscure deeper questions: Why do we accept systems that sort and rank humans at all? Who benefits from increasingly granular surveillance and prediction? What ways of being human become impossible when every action generates data to be analyzed?

[2] AI Cannot Outperform the Evidence It Was Built On

The rise of workplace AI monitoring, producing phenomena like "brain fry" documented in [« Brain Fry » : le revers caché de la révolution IA au travail], shows how quantification itself becomes a form of power. Workers report cognitive exhaustion not just from using AI tools but from being constantly measured by them—productivity scores, engagement metrics, algorithmic performance reviews. The human experience of work, with its rhythms of focus and rest, collaboration and solitude, gets flattened into data streams. Those with the power to see the dashboards, to set the metrics, to define what counts as productive, shape not just economic outcomes but the texture of daily life for millions.

The Silence of the Surveilled

Perhaps the most revealing aspect of AI discourse is not what is said but who remains unheard. While academic conferences debate algorithmic fairness and policy makers draft AI governance frameworks, those whose lives are most directly shaped by algorithmic decisions rarely appear as speakers rather than subjects. This absence is not merely an oversight—it represents a fundamental structure of power that determines whose knowledge counts as expertise and whose experiences merit consideration in shaping our technological future.

The expansion of AI surveillance in educational settings starkly illustrates these dynamics of silence. When [18] reports on children handcuffed due to algorithmic misidentification, we glimpse how surveillance systems create particular kinds of subjects: watched but not heard. The students subjected to constant monitoring, their movements tracked and behaviors analyzed, rarely participate in discussions about whether such systems should exist. Instead, their voices appear

[18] Students have been arrested for AI surveillance false alarms - NBC Chicago

only as quoted reactions to harm already done—a student describing the trauma of false arrest, a parent recounting their child's anxiety about returning to school. The power to watch is simultaneously the power to render the watched silent in conversations about watching.

This dynamic extends beyond education into every domain where algorithmic systems meet human lives. In the welfare system, those subjected to algorithmic fraud detection rarely participate in designing or evaluating these systems. The French organization La Quadrature du Net's investigation of the CAF algorithm succeeded precisely because it broke this pattern, amplifying the voices of those scored and sorted by the system. Yet such inversions remain exceptional. More commonly, as [10] documents, those affected by algorithmic welfare decisions experience them as opaque forces—denial letters citing system determinations, benefits cut by invisible calculations, lives shaped by logics they cannot access or appeal.

[10] How Do Algorithmic Decision-Making Systems Used in Public Benefits ...

The gendered dimensions of algorithmic violence reveal another layer of systematic silencing. [Deepfake, deepnude: les femmes en première ligne face aux dangers de l'intelligence artificielle] documents how women disproportionately face AI-enabled harassment, from deepfake pornography to AI-manipulated revenge porn. Yet discussions of AI safety and ethics often abstract these harms into technical challenges—how to detect synthetic media, how to authenticate images—rather than centering the voices of those targeted. The women whose images are stolen and manipulated, whose professional lives are derailed by synthetic sexual content, whose online existence becomes a battlefield, rarely shape the policy responses to these violations. Their silencing is compound: first by the technology used to violate them, then by their exclusion from conversations about preventing such violations.

Workers subject to algorithmic management face similar dynamics of enforced silence. The Uber drivers whose pay is determined by opaque algorithms, the warehouse workers whose every movement is optimized by AI systems, the call center employees whose speech patterns are analyzed for emotional authenticity—these workers generate vast amounts of data but little recognized expertise about the systems shaping their labor. When [IA au travail : 70 % des dirigeants voient des gains de productivité, les salariés restent prudents] reports the gap between executive optimism about AI and worker skepticism, it reveals how power shapes not just who benefits from AI but whose assessments of those benefits count. Workers' "prudence" gets framed as resistance to change rather than hard-won knowledge about technology's impact on their daily lives.

The geographic concentration of AI discourse amplifies these silences. Major AI conferences happen in wealthy cities; influential papers emerge from well-funded institutions; policy discussions center Global North concerns. When [1] argues for Global South AI development, she highlights how current structures ensure that AI systems designed in Silicon Valley or London get deployed across Africa without adaptation, consultation, or consideration of local contexts. The teachers in Senegal struggling with AI tutoring systems that don't recognize local languages, the farmers whose knowledge gets ignored by agricultural AI trained on industrial farming data, the communities whose ways of life become illegible to algorithmic analysis—their expertise about what doesn't work remains systematically excluded from shaping what might work better.

[1] Adji Bousso Dieng, chercheuse sénégalaise: «Les pays du Sud global doivent pouvoir développer leur propre IA»

From Diagnosis to Paralysis

The proliferation of studies documenting algorithmic bias reveals a troubling pattern: our capacity to identify problems far exceeds our ability to address them. This gap between diagnosis and action is not merely a matter of implementation challenges or technical difficulties. It represents a form of power that operates through the very act of problem identification—creating an appearance of concern while avoiding fundamental change. The 39.5% of AI discourse focused on ethical failures might seem like healthy self-criticism, but when documentation of harm fails to produce harm reduction, we must ask who benefits from endless diagnosis.

The education sector exemplifies this dynamic with painful clarity. Study after study documents bias in educational AI systems, from admissions algorithms that disadvantage minority applicants to plagiarism detectors that flag non-native English speakers at higher rates. [8] provides yet another comprehensive framework for understanding these biases, complete with taxonomies and case studies. Yet these systems continue to be deployed, often with minor tweaks that preserve their fundamental logics. The power to define a problem as "bias" rather than "discrimination," as a "technical challenge" rather than an "injustice," shapes what solutions become possible. When we frame algorithmic sorting of students as needing "debiasing" rather than questioning whether students should be algorithmically sorted at all, we've already conceded significant ground.

[8] FairAIED: Navigating Fairness, Bias, and Ethics in Educational AI ...

The response to AI detection tools in education reveals how problem identification can substitute for problem solving. Multiple studies, including [3], have demonstrated that AI detection tools are unreliable,

[3] AI Detectors Don't Work. Here's What to Do Instead.

biased against non-native speakers, and easily fooled. Yet institutions continue purchasing these tools, spending millions as documented in [21]. The existence of critical research demonstrating these tools' failures paradoxically legitimizes their continued use—institutions can claim they're aware of the limitations while still deploying systems that harm students. The market for AI detection tools thrives not despite evidence of their failures but because that evidence gets incorporated into a discourse of "ongoing improvement" and "known limitations" that normalizes their use.

[21] What Universities Spend on AI Detection — $15M+ in Data

This pattern repeats across domains. In criminal justice, studies consistently show racial bias in predictive policing algorithms. [20] provides detailed documentation of how these systems amplify existing police biases, leading to over-policing of minority communities. Yet the response is rarely to abandon predictive policing but to promise better algorithms, more training, additional oversight—solutions that preserve the fundamental power to predict and pre-empt, to treat entire communities as sites of latent criminality. The discourse of bias identification becomes a pressure release valve, allowing institutions to acknowledge problems without changing practices.

[20] UK: Police forces 'supercharging racism' with crime predicting tech ...

The workplace presents perhaps the clearest example of diagnosis without remedy. Research consistently documents how AI systems enable new forms of worker exploitation, from algorithmic wage theft to automated surveillance. Yet the solution space remains constrained to marginal adjustments—slightly better algorithms, modest regulatory tweaks, voluntary corporate commitments. When [12] warns about growing inequality between AI users and non-users, the proposed solutions focus on training programs and skill development rather than questioning why workplaces are designed to advantage those who best integrate with surveillance and automation systems. The power to define solutions within existing structures ensures that those structures remain intact.

[12] L'écart des compétences en IA se creuse : le rapport alarmant d'Anthropic révèle que les utilisateurs avancés prennent de l'avance

Even critical research can participate in this dynamic of paralysis through diagnosis. The academic tendency to end papers with calls for "further research" or "interdisciplinary collaboration" reveals how the identification of problems becomes its own form of action. When [6] meticulously documents how algorithmic systems fail vulnerable children, its recommendations focus on better documentation, improved transparency, more stakeholder engagement—all valuable but insufficient responses to systems that fundamentally transform how society relates to its most vulnerable members. The power to study harm operates differently from the power to prevent it, and those who excel at the former often lack the latter.

[6] Algorithmic Harms in Child Welfare: Uncertainties in Practice ...

The Political Economy of Algorithmic Concern

The explosion of AI ethics discourse—conferences, research centers, corporate positions—might appear as progress toward more responsible technology. Yet examining who funds this discourse, who benefits from it, and what changes it produces reveals a more complex picture. The institutionalization of AI ethics often serves to manage and contain critique rather than transform systems. This is not to dismiss the genuine concern of many working in AI ethics, but to recognize how their work operates within structures that limit its transformative potential.

Corporate investment in AI ethics reveals these dynamics most clearly. The same companies building invasive surveillance systems fund research on privacy; those creating algorithmic management tools sponsor studies on worker wellbeing; firms profiting from biased algorithms establish fairness initiatives. This is not simple hypocrisy but a form of power that operates through the appearance of self-reflection. By funding critical research, companies shape its boundaries—what questions get asked, what alternatives get explored, what recommendations seem realistic. As [4] documents, massive corporate lobbying ensures that regulatory responses remain focused on compliance procedures rather than fundamental restrictions, on transparency rather than prohibition, on mitigation rather than prevention.

[4] AI's $130M lobbying blitz hands HR the real AI compliance burden

The academic institutionalization of AI ethics follows similar patterns. Universities establish AI ethics centers often funded by the very companies whose practices merit ethical scrutiny. Research questions get shaped by fundability; career advancement depends on publication in venues that value technical contributions over social critique; interdisciplinary work remains marginalized despite rhetorical support. When [Washington enacts first AI chatbot safety law], it represents years of advocacy work translating documented harms into policy change. Yet the narrow scope of such laws—focused on specific applications rather than systemic transformation—reveals how even successful policy efforts operate within constraints set by existing power structures.

The geographic and institutional concentration of AI ethics discourse reinforces existing inequalities. Major conferences happen in expensive cities; influential positions require advanced degrees from elite institutions; recognized expertise demands fluency in technical and philosophical vocabularies remote from lived experience of algorithmic harm. When scholars from the Global South critique digital colonialism, as in [13], they must often do so in Northern academic venues, translating their critiques into forms legible to those whose in-

[13] PDF IA, colonialisme numérique et mobilisation sociale en Afrique

stitutions benefit from the very dynamics being critiqued. The power to be heard as an ethical voice operates through the same exclusionary mechanisms that AI ethics purports to challenge.

The professionalization of AI ethics creates its own dynamics of containment. Ethics review boards, responsible AI frameworks, bias audits—these mechanisms transform ethical concern into bureaucratic process. What begins as recognition of harm becomes a checklist of considerations, a set of procedures that organizations can follow to demonstrate ethical compliance without changing fundamental practices. When [14] documents rights violations in public sector AI use, the recommendations focus on procedural safeguards—impact assessments, transparency requirements, appeal mechanisms—rather than questioning whether certain uses should exist at all. The power to transform ethics into process ensures that ethical concern never threatens operational continuity.

[14] PDF Rapport algorithmes, systèmes d IA et services publics : quels droits ...

Conclusion: Breaking the Circuits of Power

The discourse around AI and social impacts has reached a critical juncture. We excel at documenting harm—the 47.3 million denied food rations, the women targeted by deepfakes, the workers surveilled and underpaid, the students arrested by algorithmic false positives. We've developed sophisticated vocabularies for discussing bias, frameworks for analyzing fairness, methodologies for measuring discrimination. Yet this very sophistication may be preventing us from asking simpler questions: Should these systems exist? Who benefits from their existence? What would it mean to refuse them?

The concentration of power in AI discourse—in who speaks, who is heard, whose concerns shape policy—mirrors and reinforces the power dynamics embedded in AI systems themselves. Those who design algorithms shape not just technical systems but social relations; those who deploy them in schools, workplaces, and welfare offices reshape how authority operates in daily life; those who study their impacts shape what counts as harm worth addressing. Meanwhile, those subjected to algorithmic decisions—watched, sorted, scored, and silenced—remain objects of study rather than subjects of power.

[ETUDE] Ce que 81 000 personnes attendent vraiment de l'IA reveals a gap between what people want from AI (augmentation, assistance, genuine help) and what they get (surveillance, assessment, control). This gap is not a bug but a feature of current power arrangements. AI systems are optimized for those who deploy them—employers seeking to minimize wages, schools seeking to detect cheat-

ing, governments seeking to reduce welfare costs—rather than those subjected to them. The discourse of AI ethics, with its focus on bias mitigation and responsible development, often accepts these applications as given, seeking to make them fairer rather than questioning their existence.

Breaking these circuits of power requires more than better algorithms or more inclusive development processes. It requires recognizing that the current explosion of AI discourse may be symptomatic of the problem rather than its solution. When documentation of harm fails to prevent harm, when identification of bias fails to eliminate bias, when ethical frameworks fail to prevent unethical applications, we must ask whether the discourse itself has become a mechanism of power—a way of managing concern without enabling change.

The path forward cannot be found in more sophisticated bias metrics or more comprehensive ethical frameworks. It requires shifting power—from those who deploy algorithmic systems to those subjected to them, from those who study impacts to those who experience them, from those who profit from AI applications to those who bear their costs. This shift cannot be achieved through discourse alone but requires material changes: workers' right to refuse algorithmic management, students' right to education without surveillance, citizens' right to access public services without algorithmic judgment.

The question is not whether AI can be made fair or ethical within existing power structures. The question is whether we can imagine and build structures where such concentrated technological power cannot accumulate in the first place. Until those whose lives are reshaped by algorithmic decisions have meaningful power to refuse those decisions, all our sophisticated discourse about AI and social impacts remains an elaborate performance—maintaining the appearance of concern while preserving the reality of control.

References

1. Adji Bousso Dieng, chercheuse sénégalaise: «Les pays du Sud global doivent pouvoir développer leur propre IA»
2. AI Cannot Outperform the Evidence It Was Built On
3. AI Detectors Don't Work. Here's What to Do Instead.
4. AI's $130M lobbying blitz hands HR the real AI compliance burden
5. Algorithmic Bias in Education | International Journal of Artificial ...

6. Algorithmic Harms in Child Welfare: Uncertainties in Practice ...

7. Anthropic Data Shows AI Skills Gap Splitting Workforces

8. FairAIED: Navigating Fairness, Bias, and Ethics in Educational AI ...

9. How companies are using AI to pay workers as little as possible

10. How Do Algorithmic Decision-Making Systems Used in Public Benefits ...

11. Inside the Suspicion Machine - WIRED

12. L'écart des compétences en IA se creuse : le rapport alarmant d'Anthropic révèle que les utilisateurs avancés prennent de l'avance

13. PDF IA, colonialisme numérique et mobilisation sociale en Afrique

14. PDF Rapport algorithmes, systèmes d IA et services publics : quels droits ...

15. Police used AI facial recognition to arrest a Tennessee woman for crimes committed in a state she says she's never visited

16. Ramificações Do Colonialismo Digital: Inteligência Artificial Como a ...

17. Scoring of welfare beneficiaries: the indecency of CAF's algorithm now ...

18. Students have been arrested for AI surveillance false alarms - NBC Chicago

19. The cultural cost of AI in Africa's education systems - UNESCO

20. UK: Police forces 'supercharging racism' with crime predicting tech ...

21. What Universities Spend on AI Detection — $15M+ in Data