The Overhead Was the Accountability: Artificial Intelligence, Human Resources, and Mission-Driven Organizations

Heather Grizzle · May 2, 2025

Small and mission-driven organizations now run enterprise HR tools without the control layer those tools were built to sit inside. What that costs, and the fix.

Contents

Enterprise human resources technology has become available to organizations that could never previously have afforded it. The governance question is what did not arrive alongside it.

What the democratization argument gets right

Enterprise human resources technology has become available to organizations that could never previously have afforded it. A twenty-person nonprofit can now run applicant ranking, structured assessment, engagement analytics, skills inference, and personalized learning delivery on a monthly subscription, using the same underlying products deployed by employers a thousand times its size. That is a real change and the enthusiasm about it is not misplaced. Capability that was concentrated in large corporations is now broadly distributed.

The argument usually advanced alongside this observation is that the distribution is not merely of capability but of advantage, and that smaller and mission-driven organizations may in fact be better positioned than corporations to use these systems well, because they are leaner, more agile, and free of the bureaucratic overhead a large human resources department carries. Novara Consulting Group's position is that this last claim inverts the actual risk, and that the inversion is the central governance problem for the sector.

The overhead is the accountability. What a large employer's human resources function contains, beneath the layers that look like inefficiency, is a set of controls: legal review before a tool is deployed, a vendor management process, records retention, a compliance officer who asks what a system measures, and a business partner whose job includes noticing when a decision looks wrong. Large corporations adopted these tools into an environment where that layer already existed. Mission-driven organizations are adopting the same tools without it, and are being told that its absence is the benefit.

The overhead is the accountability.

What did not come in the box

The software is the same. Everything around the software is different, in four respects that compound.

Nobody configured it. A small organization runs the vendor's defaults. It did not select the screening criteria, did not set thresholds, and often does not know which features are scoring candidates, because several were enabled at implementation and never reviewed. The question that eventually arrives, asking who set the criteria and on what basis, has no answer.

Nobody reviewed it. There is no counsel to ask whether an assessment measuring communication style relates to the job, no privacy function to ask where inference outputs are stored, and no procurement office to ask for validation evidence. The purchase was made by an executive director on the recommendation of a peer organization.

Nobody can negotiate it. The controls that actually protect an employer are contractual: score retention, version identification, the ability to reproduce a historical decision, access to validation evidence, cooperation in an investigation. These are settled at acquisition. A twenty-person organization signing a standard order form has no leverage to obtain them, which means the single most effective defense is the one least available at this scale.

And the legal floor does not move. Title VII and the Americans with Disabilities Act apply at fifteen employees, the Age Discrimination in Employment Act at twenty. There is no sophistication defense, no small-employer safe harbor for automated tools, and no reduction in obligation because the system was purchased rather than built. The developing case law suggests the vendor may be added as a defendant rather than substituted for one; in July 2024 the Northern District of California allowed claims against a screening software vendor to proceed on the theory that it acted as an agent of the employers using its tools. That the tool came in a box, and that nobody at the organization understood it, is not a defense. It is an admission.

The claim to strike

Of the propositions commonly advanced about artificial intelligence in mission-driven human resources, one should be removed rather than qualified: that these systems are bias-resistant, or that they reduce unconscious bias and produce fairer hiring.

There is no public evidence base supporting it, and the available evidence points the other way. Systems trained on prior successful candidates score conformity to a distribution, and the dimensions along which they most commonly score, communication, fluency, affect, and response timing, are the dimensions along which disability, accented speech, and unfamiliar background register as deviation. A complaint filed in March 2025 with the Colorado Civil Rights Division and the EEOC (Equal Employment Opportunity Commission) on behalf of a Deaf and Indigenous employee alleges precisely this pattern in an automated video interview; both companies deny the allegations and the matter is contested, but the theory of harm is now on the public record with named parties.

The audit apparatus offered as reassurance does not repair the claim. Research presented at the 2024 ACM Conference on Fairness, Accountability, and Transparency examined compliance with New York City's bias audit requirement across nearly four hundred employers and found affirmative posting rates in the low single digits, with nearly all published audits reporting impact ratios above the four-fifths threshold, a distribution that says more about who commissions audits than about the systems audited. And these instruments measure race and sex. They cannot measure disability, because employers do not know and may not lawfully ask before a conditional offer which applicants are disabled.

For a mission-driven organization this is not merely an accuracy problem. An organization that publicly commits to equitable hiring and privately relies on a vendor's fairness claim has adopted an assertion it cannot substantiate, in the one area where its credibility is its principal asset. The correct posture is that these systems are unproven on bias, that the organization is monitoring outcomes itself, and that it has not delegated the question.

Small-team analytics deanonymize by construction

The proposition that real-time engagement and performance analytics let small teams foster supportive environments deserves separate treatment, because at this scale it does something different from what it does at enterprise scale.

Engagement analytics are sold on the premise of aggregate insight: the organization sees patterns, not people. That premise depends on cell sizes large enough to obscure individuals. In an organization of twenty-five, with three teams, a dashboard reporting sentiment or engagement by team is reporting on six to nine identifiable people, and any manager who reads it knows who is who. Segment by tenure, role, or location and the reporting is individual. The anonymity is nominal, and the people being reported on generally understand this well before leadership does.

Two consequences follow. The first is that the data ceases to be reliable, because employees who correctly infer that responses are attributable answer accordingly, which is how organizations end up with dashboards showing high engagement and a resignation nobody anticipated. The second is legal. Where the analytics infer wellbeing, stress, burnout, or disengagement, the organization may be generating health-adjacent information about identifiable employees. The ADA requires medical information about employees to be maintained separately and confidentially, and an inference sitting in a general-purpose analytics product on vendor infrastructure, visible to whoever administers the platform, is unlikely to satisfy that standard. Nobody considered the question, because the system was procured as a culture tool.

The same reasoning applies to skills inference and personalized learning. A system that identifies skills gaps and targets development is producing a profile that determines who receives investment. Allocating development on the basis of an inferred profile is an employment decision, and it is reviewable as one. If the inference draws on tenure, role history, or engagement signals correlated with age or disability, the organization has built an exposure and recorded it as a benefit.

What a lean organization should actually keep

The recommendation is not a compliance program, which an organization of this size cannot operate and should not attempt. It is a small set of controls that survive being run by one person with other responsibilities. They are cheap at this scale, and they are the part of the enterprise overhead worth importing.

Know what is running. List every point where software evaluates, ranks, filters, scores, or profiles a person, including features enabled by default at implementation. Most organizations doing this for the first time find at least one system nobody knew was scoring.

Turn off what cannot be explained. Where nobody can state what a feature measures and why it matters to the job, disable it rather than document it. This option is genuinely available to small organizations and largely unavailable to large ones, and it is the highest-yield action on this list.

Ask the vendor in writing, and keep the answer. What does this feature score. What does it infer, from which signals, and where is the inference stored. Do you retain scores and configuration, for how long, and can you reproduce a decision from eighteen months ago. What validation evidence exists. What accommodation pathway does the product support. A refusal is also an answer and is worth retaining.

Build the accommodation pathway before it is needed. Name who decides, set a turnaround, ensure an alternative assessment exists that does not disadvantage whoever uses it, and state its availability on the assessment itself rather than on a careers page. Keep the file of requests and outcomes. This is the cheapest control available and it addresses the claim most likely to arrive.

Set a floor for any people analytics. No reporting on groups below a stated minimum size, no individual attribution, and no inference about wellbeing or health retained in a general-purpose system.

Ask for retention and reproduction terms at renewal. The request may be refused. Asking and being refused is materially better evidence of institutional diligence than a memorandum written after a charge is filed.

THE SIX CONTROLS, IN SHORT

Know what is running; turn off what cannot be explained; ask the vendor in writing and keep the answer; build the accommodation pathway before it is needed; set a floor for any people analytics; ask for retention and reproduction terms at renewal.

Conclusion

The distribution of enterprise capability to organizations that could not previously afford it is a genuine gain, and the sector should take it. What has not been distributed is the apparatus that made those systems governable inside the organizations they were built for, and the language currently used to describe its absence, agility, leanness, freedom from overhead, obscures what is missing.

A twenty-person organization deploying automated screening holds the same obligations as a twenty-thousand-person one, with fewer resources to meet them and, if it is mission-driven, considerably more to lose when it cannot. The controls that close the gap are not expensive at this scale. They are a list, a set of questions, a written file, and a decision about what to switch off. The organizations that install them will not merely be safer. They will know what their systems are doing, which is the thing everyone else in this conversation is assuming and almost nobody has verified.

References

- American Civil Liberties Union, ACLU of Colorado, Public Justice, and Eisenberg & Baum LLP. Complaint filed with the Colorado Civil Rights Division and the U.S. Equal Employment Opportunity Commission, March 19, 2025.
- Americans with Disabilities Act of 1990, 42 U.S.C. § 12112, including § 12112(d) on medical examinations, disability-related inquiries, and confidentiality; Title VII of the Civil Rights Act of 1964, 42 U.S.C. § 2000e; Age Discrimination in Employment Act, 29 U.S.C. § 621.
- Mobley v. Workday, Inc., No. 3:23-cv-00770 (N.D. Cal.), order on motion to dismiss, July 2024.
- New York City Local Law 144 of 2021, Automated Employment Decision Tools.
- Uniform Guidelines on Employee Selection Procedures, 29 C.F.R. Part 1607; EEOC recordkeeping requirements, 29 C.F.R. Part 1602.
- Wright, Lucas, et al. "Null Compliance: NYC Local Law 144 and the Challenges of Algorithm Accountability." Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency.