NOVARA
CONSULTING GROUP

This document was translated from English by machine and may contain mistakes. The English original is at https://www.novaracg.com/2025/05/02/the-overhead-was-the-accountability-artificial-intelligence-human-resources-and-mission-driven-organizations/.

管理成本本身即是问责机制:人工智能、人力资源与使命驱动型组织

Heather Grizzle · May 2, 2025

小型且以使命为导向的组织如今在运行企业级人力资源工具，却缺少这些工具本应嵌入其中的管控层。这会带来怎样的代价，以及解决之道。

目录

1. “民主化”论点说对了什么
2. 盒子里没有附带的东西
3. 应当摒弃的说法
4. 小团队分析在结构上就会“去匿名化”
5. 精简型组织真正应当保留的东西
6. 结语
7. 参考文献

企业级人力资源技术已经普及到那些以前根本负担不起它的组织手中。治理层面的问题在于,伴随它一同到来的东西并没有到来。

“民主化”论点说对了什么

企业级人力资源技术已经普及到那些以前根本负担不起它的组织手中。一个二十人的非营利组织如今可以以每月订阅的形式运行申请人排名、结构化评估、参与度分析、技能推断和个性化学习交付,使用的是与规模是其千倍的雇主所部署的相同底层产品。这是一个真实的变化,人们对此的热情并非没有道理。曾经集中在大型企业手中的能力,如今已被广泛分散。

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.

> 开支就是问责。

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.

盒子里没有附带的东西

软件是一样的。软件周围的一切都不同,体现在四个相互叠加的方面。

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.

没有人审查过它。没有法律顾问去追问衡量沟通风格的评估是否与工作相关,没有隐私职能部门去追问推断结果存储在哪里,也没有采购部门去索要验证证据。这项采购是由执行主任根据同行组织的推荐而做出的。

没有人能就它进行谈判。真正保护雇主的控制机制是合同性质的:分数保留、版本识别、还原历史决策的能力、获取验证证据的权限、在调查中的配合义务。这些都是在采购时确定的。一个签署标准订购单的二十人组织没有任何筹码去争取这些条款,这意味着最有效的那道防线,恰恰是在这种规模下最难获得的。

而法律底线并不会因此移动。《第七章》(Title VII)和《美国残疾人法案》适用于雇员达十五人的企业,《雇佣年龄歧视法》适用于雇员达二十人的企业。不存在“复杂程度”这样的抗辩理由,也没有针对自动化工具的小雇主安全港条款,更不会因为系统是购买而非自行开发就减轻义务。正在发展中的判例法表明,供应商可能被追加为被告,而不是取代雇主成为被告;2024年7月,加州北区法院允许针对一家筛选软件供应商的诉讼请求继续进行,理由是该供应商充当了使用其工具的雇主的代理人。工具是现成打包购买的、组织里没有人真正理解它,这并不能成为抗辩理由,反而是一种自我认罪。

应当摒弃的说法

在人们通常提出的关于人工智能在以使命为导向的人力资源领域应用的各种说法中,有一种应当被彻底摒弃,而非有所保留地修正:即这些系统能够抵抗偏见,或者说它们能够减少无意识偏见、产生更公平的招聘结果。

没有公开的证据基础支持这一说法,而现有的证据恰恰指向相反的方向。基于以往成功候选人训练出来的系统,衡量的是对某种分布的符合程度,而它们最常用来打分的那些维度——沟通、流利度、情感表达和回应时长——恰恰是残障人士、带口音者和背景不熟悉者最容易被判定为“偏离”的那些维度。2025年3月,一名聋人及原住民员工向科罗拉多州民权司(Colorado Civil Rights Division)和EEOC(平等就业机会委员会)提起的一项投诉,正是针对一次自动化视频面试提出了这种模式的指控;涉案两家公司均否认这些指控,案件仍在争议之中,但这种伤害理论如今已经带着具名当事方进入了公共记录。

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.

小团队分析在结构上就会“去匿名化”

“实时参与度和绩效分析能让小团队营造支持性环境”这一说法应当单独讨论,因为在这种规模下,它所起的作用与在企业规模下完全不同。

参与度分析的售卖前提是聚合洞察:组织看到的是模式,而非具体的人。这一前提依赖于足够大的样本单元规模,以掩盖个体身份。在一个只有二十五人、分为三个团队的组织中,一份按团队报告情绪或参与度的仪表盘,实际上报告的是六到九个可识别的人,任何阅读它的管理者都知道谁是谁。如果再按任职年限、职位或地点细分,报告就变成了针对个人的。所谓的匿名性只是名义上的,而被报告对象的人通常比管理层更早、更清楚地意识到这一点。

由此产生两个后果。第一,数据不再可靠,因为员工一旦正确推断出回答是可以被追溯到个人的,就会相应地作答,这正是为何组织最终会出现仪表盘显示参与度很高、却发生了无人预料的离职事件。第二个后果是法律层面的。当分析结果推断出幸福感、压力、职业倦怠或脱离投入等信息时,组织可能正在生成关于可识别员工的健康相关信息。《美国残疾人法案》要求员工的医疗信息必须单独、保密地保存,而一个存放在供应商基础设施上的通用分析产品中、对该平台管理人员可见的推断结果,不太可能满足这一标准。没有人考虑过这个问题,因为这个系统当初是作为一种“文化工具”被采购的。

同样的道理也适用于技能推断和个性化学习。一个识别技能差距并针对性地提供发展机会的系统,实际上生成了一份决定谁能获得投资的档案。基于推断出的档案来分配发展资源,是一项雇佣决策,理应作为雇佣决策接受审查。如果这种推断依据的是任职年限、职位履历或与年龄、残障相关的参与度信号,那么这个组织就是在制造一种风险敞口,却把它记录成了一项福利。

精简型组织真正应当保留的东西

这里给出的建议不是一套合规项目——这种规模的组织既无法运作,也不应尝试。它是一小组控制措施,即便由一个身兼数职的人来执行也能维持下去。在这种规模下,它们的成本很低,也是企业级开支中真正值得借鉴的部分。

了解正在运行的是什么。列出每一个软件对人进行评估、排名、筛选、打分或建档的环节,包括那些在实施时默认启用的功能。大多数第一次这样做的组织都会发现,至少有一个系统在打分,而此前没有人知道。

关闭那些无法解释的功能。如果没有人能说清楚某项功能衡量的是什么、以及它与该岗位有何关系,就应当关闭它,而不是把它记录下来。这个选项对小型组织来说是真正可行的,而对大型组织来说则基本不可行,这也是本清单中收益最高的一项行动。

以书面形式询问供应商,并保留答复。这项功能衡量的是什么?它推断的是什么、依据哪些信号、推断结果存储在哪里?你们是否保留分数和配置信息、保留多久、能否还原十八个月前的一项决策?存在哪些验证证据?该产品支持哪种便利安排(accommodation)途径?拒绝回答本身也是一种答案,同样值得保留。

在需要之前就建立便利安排途径。指定由谁来决定,设定一个处理时限,确保存在一种不会使使用者处于不利地位的替代评估方式,并在评估本身而非招聘页面上说明其可用性。保留申请与结果的记录档案。这是可获得的成本最低的控制措施,也是应对最可能出现的指控的手段。

为任何"人员分析"设定一条底线。不对低于规定最小规模的群体进行报告,不进行个人归因,不在通用系统中保留关于幸福感或健康状况的推断信息。

在续约时询问数据保留和还原条款。这一请求可能会被拒绝。提出请求并被拒绝,作为机构尽职调查的证据,要比在指控提出之后才撰写的备忘录有力得多。

> **六项控制措施,简而言之**
>
> 了解正在运行的是什么;关闭那些无法解释的功能;以书面形式询问供应商并保留答复;在需要之前就建立便利安排途径;为任何人员分析设定一条底线;在续约时询问数据保留和还原条款。

结语

企业级能力向此前负担不起它的组织扩散,是一项真实的收获,该行业理应接受它。尚未被扩散开来的,是那套使这些系统在其最初设计所服务的组织内部变得可治理的机制,而目前用来描述这种缺失的说法——灵活性、精简、免于开支负担——恰恰掩盖了缺失的是什么。



一个部署自动化筛选系统的二十人组织,承担着与一个两万人组织相同的义务,却拥有更少的资源去履行这些义务,而如果它是以使命为导向的,一旦无法履行,它将失去的也要多得多。弥合这一差距的控制措施在这种规模下并不昂贵。它们只是一份清单、一组问题、一份书面档案,以及一个关于关闭什么的决定。安装这些措施的组织不仅会更加安全,还会真正了解自己的系统在做什么——而这正是这场讨论中其他所有人都在想当然认为成立、却几乎无人真正核实过的事情。

参考文献

- 美国公民自由联盟、科罗拉多州公民自由联盟、公共正义组织以及Eisenberg & Baum LLP律师事务所。向科罗拉多州民权司及美国平等就业机会委员会提交的投诉,2025年3月19日。
- 1990年《美国残疾人法案》,42 U.S.C. § 12112,包括涉及医疗检查、与残障相关询问及保密性的 § 12112(d);1964年《民权法案》第七章,42 U.S.C. § 2000e;《雇佣年龄歧视法》,29 U.S.C. § 621。
- Mobley诉Workday公司案,案号3:23-cv-00770(加利福尼亚北区联邦地区法院),2024年7月驳回动议裁定。
- 《2021年纽约市地方法144号》，自动化就业决策工具。
- 《雇员甄选程序统一指南》,29 C.F.R. Part 1607;EEOC 记录保存要求,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.

TIPS

Know something about a sign language AI system that the public should know? Send it to tips@novaracg.com. Every tip is read by me and kept confidential. If you work in the UK, writing to us is not by itself a protected disclosure under the Public Interest Disclosure Act 1998, so before you share employer information, check your options with Protect (protect-advice.org.uk), the free whistleblowing advice line.

引用方式

Grizzle, H. M. (2025, May 2). 管理成本本身即是问责机制:人工智能、人力资源与使命驱动型组织. Novara Consulting Group. https://www.novaracg.com/zh/2025/05/02/the-overhead-was-the-accountability-artificial-intelligence-human-resources-and-mission-driven-organizations/