NOVARA
CONSULTING GROUP

Instruments, Not Principles: Sequencing AI Readiness in Human Resource Governance

Heather Grizzle · June 7, 2025

Most AI readiness models put the ethics layer third. By then the systems are in production, procurement has closed, and governance is retrospective.

Contents

1. A readiness model that arrives too late
2. Why principles are not controls
3. The readiness sequence, inverted
4. Talent intelligence as the illustrative case
5. What this changes about the consulting question
6. References

A readiness model that arrives too late

Most organizational AI readiness frameworks share a structural feature that undermines them. They sequence capability before control. Data foundations are established first, applications are deployed second, and an ethics or trust layer is introduced third, typically described in terms of guidelines, principles, and commitments to fairness and transparency. The sequence is intuitive, because it mirrors how the work actually gets funded, and it is the sequence most readiness models in circulation recommend.

It is also the sequence that produces the failures now visible in employment AI. By the time an ethics layer is designed, the systems it would govern are in production, procurement has closed, contractual terms are fixed, vendor selection is sunk, and the organization's practical options have narrowed to documentation and disclosure. Governance introduced at step three is not governance. It is retrospective justification of decisions made at step two.

The correction is not to add an ethics stage earlier. It is to recognize that ethical guidelines are not the instrument the situation requires, and that treating them as one is itself the error.

> Governance introduced at step three is not governance.

Why principles are not controls

The distinction between a principle and an instrument is not semantic, and it has been examined carefully enough in the governance literature that organizations need not rediscover it.

Salamon′s analysis of the tools of government establishes that policy objectives are realized through instruments with identifiable properties: what activity they authorize, who administers them, what compliance they compel, and what they make observable. A stated commitment does none of these things. It expresses an intention without specifying an actor, a trigger, an evidentiary output, or a consequence for non-performance. Applied to artificial intelligence, the pattern has been documented directly. Jobin, Ienca, and Vayena′s survey of the global landscape of AI ethics guidelines found substantial convergence on high-level principles alongside marked divergence in how those principles were interpreted and what they were understood to require. Mittelstadt′s assessment reached the sharper conclusion that principles alone cannot guarantee ethical artificial intelligence, in the absence of the professional structures, accountability mechanisms, and enforcement pathways that make principles operative in other domains.

An organization that responds to employment AI risk by developing clear ethical guidelines has therefore adopted the intervention with the weakest evidentiary record. This is not because fairness and transparency are unworthy objectives. It is because a guideline produces no artifact, and every obligation the organization actually faces is discharged through artifacts.

Consider the commitment to transparency in how algorithms work, which appears in nearly every readiness model. It is not achievable on the terms stated. Vendors will not disclose model architecture, most systems in this category are not interpretable in the sense the commitment implies, and an organization that has promised algorithmic transparency has promised something it cannot deliver and will be asked about. The

achievable and more useful commitment is different in kind: documented evidence of outcomes across protected classes, produced on a stated cadence, by a party the organization can identify. That is an instrument. It specifies who does what, when, and what record results.

The readiness sequence, inverted

A readiness model built on instruments rather than principles orders the work differently, and the reordering is the substantive contribution rather than a presentational change.

Coverage determination comes first. Before any system is evaluated on capability, the organization must be able to answer whether the system's outputs are used to make, guide, or materially influence decisions about hiring, promotion, compensation, assignment, discipline, or continued employment. That formulation, and not the product category under which a system is marketed, determines what obligations attach across every regime currently governing employment AI. Coverage determination is an instrument because it produces a record, assigns an owner, and triggers a defined workflow. It is also cheap, and it is the only step that becomes more expensive the longer it is deferred.

Data governance follows, reframed. Readiness models describe the data objective as a single source of truth, which is a data quality goal. The governance objective is different and narrower: the organization must be able to establish what data a given inference was drawn from, when, and with what provenance. Clean integrated data supports better predictions. Documented data provenance supports the ability to answer a regulator or a plaintiff about why a person received the outcome they received. Only the second is a governance capability, and an organization can possess the first entirely without it.

Evidentiary requirements enter at procurement, not after. This is the point at which the organization retains leverage. Bias audit documentation identifying methodology, tested classes, and population. Confirmation that no geographic, educational, or tenure-derived feature functions as a proxy. Specification of the notice and adverse action workflow and of which party performs it. Retention and export terms sufficient to survive platform migration across the applicable limitations period. Each of these is negotiable before award and is not negotiable afterward, which is the entire argument for the sequencing.



Deployment carries monitoring, not affirmation. The commitment that matters post-implementation is outcome testing at a stated interval, because under a disparate impact standard an organization's belief that its tool is fair is not responsive to the question of whether outcomes differ. Only measurement is responsive.

The record layer closes the loop. Attestation bound to individual, version, date, and authenticated pathway. Retention architected at creation. The organization's ability to demonstrate what it did is the output of the whole sequence, and it is the only output an adverse party will ever see.

The model is a cycle rather than a ladder because the first stage recurs. Systems enter organizations continuously, frequently through functions that do not classify them as employment decision technology, and coverage determination is therefore a monitored condition rather than a completed project.

Talent intelligence as the illustrative case

The readiness models in circulation typically place proactive talent intelligence, mapping existing skills against future requirements and generating data-driven career paths, at the second stage, ahead of the governance layer. The placement is instructive because that specific capability is among the most heavily regulated in the entire employment AI category, and is almost never recognized as such.

A system that maps employee capability and generates career path recommendations is producing assessments that materially influence promotion and assignment. New York City's Local Law 144 reaches promotion decisions. The Illinois amendment to the Human Rights Act, effective January 2026, reaches promotion, discipline, and discharge, and imposes notice obligations attaching to use rather than to outcome. Annex III of the EU Artificial Intelligence Act enumerates promotion, task allocation, and performance evaluation by function rather than by product category. An organization that deploys talent intelligence at stage two and reaches its trust framework at stage three has deployed a regulated system and then written its principles about it.

This is the practical cost of the conventional sequence, and it is not hypothetical. It is the ordinary path by which organizations acquire obligations they did not evaluate.

What this changes about the consulting question

The question organizations ask is which AI tools they should adopt. The question that determines their exposure is which of the systems they have already adopted are consequential decision systems, and what evidence they hold about each.

Those are different questions, and the second is answerable immediately, without a technology decision, at low cost, and with a durable artifact as the output. It is also the question that makes the subsequent technology decisions defensible rather than merely enthusiastic. An organization that can state, for every system touching an employment decision, what it does, what obligations attach, what evidence exists, and who owns the next review, has built something. An organization with a set of ethical AI principles has written something.

Integrating artificial intelligence into human resource operations is not a technology project. That much the conventional models get right. It is a governance project, and governance is constituted by instruments, not by commitments.

References

- Salamon, L. M., ed. *The Tools of Government: A Guide to the New Governance*. Oxford University Press, 2002.
- Jobin, A., Ienca, M., and Vayena, E. "The global landscape of AI ethics guidelines." *Nature Machine Intelligence* 1 (2019): 389–399.
- Mittelstadt, B. "Principles alone cannot guarantee ethical AI." *Nature Machine Intelligence* 1 (2019): 501–507.
- National Institute of Standards and Technology, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, January 2023.
- ISO/IEC 42001:2023, *Information technology — Artificial intelligence — Management system*.
- New York City Administrative Code § 20-870 et seq. (Local Law 144 of 2021).
- Illinois House Bill 3773, 103rd General Assembly (2024), amending the Illinois Human Rights Act, effective 1 January 2026.
- Regulation (EU) 2024/1689, Annex III(4).
- Novara Consulting Group, "The Development Exception: Internal Mobility Platforms and the Perimeter of Employment AI Governance," May 2025.

* * *

Disclaimer.
Novara Consulting Group provides governance and procurement advisory services for institutional artificial intelligence deployment and workforce compliance infrastructure. This article is analytical and does not constitute legal advice.

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.

How to cite

Grizzle, H. M. (2025, June 7). Instruments, Not Principles: Sequencing AI Readiness in Human Resource Governance. Novara Consulting Group. https://www.novaracg.com/2025/06/07/instruments-not-principles-sequencing-ai-readiness-in-human-resource-governance/