Research Article

Decision Support or High-Risk? Classifying AI Workforce-Intelligence Systems Under the EU AI Act

Sathyaraj Asaithambi[1]

1. Independent researcher

The EU Artificial Intelligence Act (Regulation (EU) 2024/1689) designates AI systems used in employment, worker management, and access to self-employment as high-risk under Annex III. Such systems attract stringent obligations for risk management, data governance, transparency, human oversight, record-keeping, and conformity assessment. A fast-growing class of enterprise systems sits ambiguously against this category: external workforce-intelligence platforms that convert public labour-market signals, such as job postings, competitor movements, and skill-demand shifts, into aggregate strategic guidance. These platforms inform workforce planning at the level of roles, functions, and populations rather than making decisions about identifiable individuals. The boundary between high-risk employment AI and governed strategic decision support is, however, neither self-evident nor consistently applied in practice.

This paper develops a classification framework that distinguishes workforce-intelligence architectures falling within the Act's high-risk employment scope from those that constitute defensible decision support. Following the design-science research paradigm, and drawing on the derogation criteria in Article 6(3), it analyses four system archetypes and identifies five architectural properties that materially affect classification and defensibility: an aggregation boundary, deterministic and reproducible processing, explain-only outputs, human-in-the-loop escalation, and end-to-end provenance logging. The properties are instantiated on a generalised reference architecture, translated into an operational compliance checklist, and extended into a practice layer that covers vendor and product assurance, effectiveness and bias testing across the lifecycle, post-deployment monitoring and re-classification triggers, and an operating model of roles with worker-representative consultation. The paper offers HR leaders, system designers, and governance functions

 doi.org/10.32388/YXJXoT

a practical instrument for the Act's phased enforcement, and sets out a research agenda for responsible workforce intelligence.

Correspondence: papers@team.qeios.com — Qeios will forward to the authors

1. Introduction

The European Union's Artificial Intelligence Act entered into force on 1 August 2024 and applies in phases. Its prohibitions on certain practices took effect on 2 February 2025, its obligations for general-purpose AI models on 2 August 2025, and the bulk of the high-risk regime becomes applicable on 2 August 2026. The Act adopts a risk-based structure. Among the standalone systems it treats as high-risk are those used in employment and worker management, enumerated in Annex III. For any organisation deploying AI in the flow of talent decisions, the compliance stakes are now concrete rather than prospective.

A distinct class of enterprise system has, at the same time, moved from novelty to normal. Workforce-intelligence platforms ingest public and aggregate labour-market signals such as job postings, competitor hiring patterns, executive and leadership movements, emerging skill demand, geographic hub formation, and disclosed restructurings. They convert those signals into forward-looking, population-level guidance for strategic workforce planning. Such systems increasingly sit alongside classical people analytics in the operating model of talent, competitive-intelligence, and workforce-planning functions. They are built to answer questions such as which capabilities an organisation should build, where a function is most exposed to automation, and where scarce roles are migrating across the market.

These two developments intersect awkwardly. Annex III's employment category is framed around AI that recruits, selects, evaluates, or manages identifiable natural persons. Workforce-intelligence systems of the kind just described are typically designed to operate on aggregate signals and to produce guidance about roles and populations, not verdicts about individuals. Whether such systems fall inside the high-risk employment category, or instead constitute governed strategic decision support outside it, is not obvious. The intuitive answer that 'it is only aggregate data, so it cannot be high-risk' is, as this paper shows, insufficient. Classification under the Act turns on architecture, deployment discipline, and use, not on data type alone.

The existing literature does not resolve the question. A substantial body of legal scholarship analyses the Act's structure and significance in general terms, and sectoral translations of the high-risk regime have begun to appear, notably in health care. Separately, philosophical and information-systems work has begun to formalise how AI systems can be classified for governance purposes; the most direct treatment proposes complementary mental models (the switch, the ladder, and the matrix) that give organisations a vocabulary for demarcating the scope of AI governance frameworks[1]. Information-systems research has, in parallel, produced organisation-level frameworks for responsible AI governance. The most comprehensive recent synthesis defines responsible AI governance through structural, relational, and procedural practices, and calls for work on how such principles are operationalised[2]. What is missing is a treatment pitched at the level of the system itself: a classification framework for external workforce-intelligence architectures, and a specification of the design choices that place a given system on one or the other side of the high-risk line.

Stated as a design-science research question[3][4], the paper asks: how should an external workforce-intelligence system be built, procured, tested, monitored, and governed so that its classification under the EU AI Act is clear and defensible at deployment, and remains defensible through operation? The artefacts produced, a taxonomy, five architectural properties, a reference architecture, an operational checklist, and a practice layer for assurance and governance, form the paper's answer. Their utility is judged against the Act's criteria and against the working needs of practitioners in HR, workforce-planning, engineering, legal, and governance functions.

This paper makes four contributions. First, it offers a taxonomy that separates employee-facing people analytics from external workforce intelligence, and distinguishes four system archetypes by data subject and decision proximity. Second, it maps those archetypes onto the Act's risk tiers, giving particular weight to the derogation in Article 6(3). That article allows an Annex III system to be treated as not high-risk where it does not pose a significant risk to health, safety, or fundamental rights, subject to the important proviso that a system is always high-risk where it profiles natural persons. Third, it identifies five architectural properties that materially affect classification and defensibility, instantiates them on a generalised reference architecture, and expresses them as an operational compliance checklist that practitioners can apply directly. Fourth, it extends the framework into a practice layer that covers vendor and product assurance, effectiveness and bias testing across the lifecycle, post-deployment monitoring and re-classification triggers, and an operating model of roles that make and sustain the classification.

The design-science method organises the paper as follows, following Peffers et al.[4]. Problem identification and motivation are set out in Sections 1 and 2 and the ambiguity described above. Objectives of the solution are the classification and defensibility criteria drawn from Article 6(3) in Section 2. Design and development produce five artefacts, in Sections 3, 5, 6, and 7: the archetype taxonomy, the five architectural properties, the reference architecture, the compliance checklist, and the practice layer. Demonstration is provided through the worked illustration in Section 6 and the vendor-assurance and RACI applications in Section 7. Evaluation is scoped for future empirical work in Section 8, following the Framework for Evaluation in Design Science[5]. Communication is the present working paper.

The analysis is scoped as design-oriented rather than doctrinal. It is written for the people who design, procure, and govern these systems, and it treats the classification question as a design problem: given the Act's criteria, how should a workforce-intelligence system be built and operated so that its regulatory position is clear and defensible, and so that residual risk is controlled if the system turns out to be high-risk after all? The remainder proceeds as follows. Section 2 sets out the Act's risk architecture and its employment category. Section 3 defines workforce intelligence and presents the archetype taxonomy. Section 4 works each archetype through the classification criteria and surfaces the genuine ambiguities. Section 5 develops the five architectural properties. Section 6 presents the reference architecture and checklist. Section 7 extends the framework into a practice layer for vendor assurance, effectiveness and bias testing, monitoring, and the operating model. Section 8 draws out implications and a research agenda. Section 9 concludes.

2. The EU AI Act's risk architecture and the employment category

Regulation (EU) 2024/1689 establishes a horizontal, risk-based regime that applies across sectors and, through its market-placing logic, reaches providers and deployers outside the Union whose systems are used within it. It sorts AI systems into tiers by the level of risk they present. A small set of practices is prohibited outright under Article 5. A limited set of systems is subject only to transparency obligations under Article 50, for example the duty to disclose that a person is interacting with an AI system or that content has been artificially generated. The great majority of systems fall into a minimal-risk residual category that the Act does not specifically regulate. Between the prohibited and the minimal sits the high-risk tier, which carries the Act's most demanding obligations and is the tier that matters for workforce systems.

High-risk status arises in two ways. Under Article 6(1), a system is high-risk if it is a safety component of, or is itself, a product covered by the Union harmonisation legislation listed in Annex I. Under Article 6(2), a system is high-risk if it is intended to be used for one of the purposes enumerated in Annex III. Annex III lists eight areas: biometrics; critical infrastructure; education and vocational training; employment, workers management and access to self-employment; access to essential private and public services; law enforcement; migration, asylum and border control; and the administration of justice and democratic processes.

2.1. The employment category

Point 4 of Annex III concerns employment, workers management, and access to self-employment. It reaches, first, AI systems intended to be used for the recruitment or selection of natural persons, in particular to place targeted job advertisements, to analyse and filter job applications, and to evaluate candidates. It reaches, second, AI systems intended to be used to make decisions affecting the terms of work-related relationships, the promotion or termination of such relationships, the allocation of tasks based on individual behaviour or personal traits, or the monitoring and evaluation of the performance and behaviour of persons in such relationships. Two features of this drafting are decisive for what follows. The category is framed throughout around natural persons: candidates, applicants, and workers within a relationship. It is also framed around decisions and evaluations directed at those persons. A system that neither concerns identifiable natural persons nor produces decisions or evaluations about them does not obviously answer to this language.

2.2. What high-risk status entails

If a system is high-risk, a substantial compliance apparatus follows. Providers must establish a risk-management system (Article 9), meet data and data-governance requirements (Article 10), maintain technical documentation (Article 11), keep records through automatic logging (Articles 12 and 19), ensure transparency and provide information to deployers (Article 13), enable effective human oversight (Article 14), and achieve appropriate accuracy, robustness, and cybersecurity (Article 15). Deployers carry their own obligations under Article 26, and in defined circumstances must complete a fundamental-rights impact assessment under Article 27. High-risk systems in the Annex III areas must be registered in the EU database under Article 49. These obligations are cumulative and evidentiary: compliance is demonstrated through documentation, not asserted.

2.3. The Article 6(3) derogation and the profiling proviso

Annex III listing is not conclusive. Article 6(3) provides that a system referred to in Annex III is not to be considered high-risk where it does not pose a significant risk of harm to the health, safety, or fundamental rights of natural persons, including by not materially influencing the outcome of decision-making. The Act specifies conditions under which this holds: where the system performs a narrow procedural task; where it improves the result of a previously completed human activity; where it detects decision-making patterns or deviations from prior patterns and is not meant to replace or influence a previously completed human assessment without proper human review; or where it performs a preparatory task to an assessment relevant to the listed use cases. This derogation is the hinge on which the decision-support question turns, because a workforce-intelligence system that advises rather than decides may fit its logic.

The derogation carries a hard limit. Notwithstanding those conditions, a system referred to in Annex III is always high-risk where it performs profiling of natural persons. This proviso operates as a bright line. Any drift of a workforce-intelligence system toward inference about identifiable individuals collapses the derogation and returns the system to the high-risk tier. A provider that considers an Annex III system not to be high-risk must, under Article 6(4), document that assessment before the system is placed on the market or put into service, and remains subject to the registration obligation in Article 49(2). Article 80 sets out the procedure by which authorities may contest such a self-classification. Articles 6(6) and 6(7) further empower the Commission to adopt delegated acts amending or adding to the derogation conditions, so the criteria discussed here are best treated as a stable but evolving reference rather than a fixed test. The Act also anticipates that the Commission will issue guidelines with practical examples distinguishing high-risk from non-high-risk use cases, which will refine the picture over time. Classification is therefore a documented, contestable judgement rather than a label that attaches automatically.

3. What 'workforce intelligence' is: a taxonomy

People analytics, sometimes called HR analytics, is conventionally understood as the use of workforce data and analytical techniques to improve HR and management decisions. In its established form it is internal and largely employee-level. It draws on human-resource information systems and related records to describe, and increasingly to predict, phenomena such as attrition, engagement, performance, and mobility for identifiable employees or small teams. Its promise and its pitfalls have been debated for

a decade, including sober assessments of whether the field has matured into a strategic capability[6]. Whatever its trajectory, its centre of gravity is the internal workforce and, frequently, the individual within it.

Workforce intelligence, as used here, has a different centre of gravity. It is external, aggregate, and forward-looking. It reads signals from the labour market rather than from the internal HRIS: the volume and composition of job postings, the hiring and reorganisation moves of competitors, disclosed leadership changes, the emergence and decay of skills, and the geographic redistribution of roles. Its outputs are pitched at the level of roles, functions, markets, and scenarios: where capability is scarce, which functions are most exposed to automation, how a talent hub is forming. It is closer in spirit to competitive intelligence and strategic planning than to case-level HR administration.

The distinction that matters for the Act is not internal-versus-external as such, but two orthogonal dimensions: the data subject the system operates on, and the proximity of its outputs to a decision about a specific person. On the first dimension, a system may operate on identifiable natural persons or on aggregate entities such as roles, populations, and markets. On the second, a system may produce or materially influence a decision about a specific worker, or it may inform strategy at a level that does not resolve to any individual. Crossing these dimensions yields four archetypes, summarised in Table 1.

Archetype	What it does	Data subject / output level	Likely AI Act position
A. Individual scoring	Screens, ranks, or scores identifiable candidates or employees; drives or feeds a hiring, promotion, or exit decision.	Identifiable natural persons; individual-level output.	High-risk. Squarely within Annex III(4)(a)/(b); profiling carve-out engaged.
B. Internal aggregate analytics	Team- or unit-level attrition, engagement, or capability analytics from internal HRIS data.	Aggregate, but derived from identifiable employees.	Context-dependent. High-risk if it profiles persons or materially influences individual decisions; otherwise arguable under Art. 6(3).
C. External talent-market intelligence	Reads public labour-market signals: skill demand, competitor hiring, leadership moves, hub formation.	Aggregate roles, markets, populations; no identifiable individuals.	Plausibly outside Annex III(4) or within the Art. 6(3) derogation, if architecture holds.
D. Workforce planning / AI-exposure modelling	Models task-level exposure, scenario forecasts, and ranked strategic actions at role / function level.	Roles, functions, scenarios; population-level output.	Plausibly decision support under Art. 6(3), provided outputs do not resolve to individuals.

Table 1. Four archetypes of workforce analytics and intelligence.

Archetype A is individual scoring: systems that screen, rank, or evaluate identifiable candidates or employees and feed a hiring, promotion, task-allocation, or exit decision. Archetype B is internal aggregate analytics: team- or unit-level analytics derived from internal records, aggregate in output but grounded in identifiable persons. Archetype C is external talent-market intelligence: systems that sense public labour-market signals and report at the level of skills, markets, and competitor behaviour. Archetype D is workforce planning and AI-exposure modelling: systems that decompose roles into tasks, model exposure and scenarios, and produce ranked strategic actions at the level of roles and functions. Archetypes C and D are the under-analysed class, and the balance of this paper concentrates on them.

4. The classification problem

Working the archetypes through Article 6 and Annex III shows why the aggregate-data intuition is inadequate and where the genuine difficulty lies.

Archetype A. Individual scoring is the paradigm case of the employment category. A system that analyses and filters applications or evaluates candidates falls within Annex III point 4(a); a system that informs promotion, termination, task allocation, or performance evaluation falls within point 4(b). Because such systems operate on identifiable natural persons and typically profile them, both the Annex III listing and the profiling proviso are engaged, and the Article 6(3) derogation is unavailable. These systems are high-risk, and the analysis is straightforward.

Archetype B. Internal aggregate analytics is context-dependent. Where team-level output is derived by profiling identifiable employees, or where the analytics materially influence decisions about particular workers (for example by flagging individuals or small groups for intervention), the profiling proviso and the 'materially influencing the outcome' test point toward high-risk. Where the analytics are truly aggregated, do not profile individuals, and are not used to replace or influence a human assessment of any particular person without proper review, an Article 6(3) argument becomes available. The determining factors are whether individuals are profiled and whether outputs influence individual decisions, not the mere fact that a number is reported at team level.

Archetypes C and D. External talent-market intelligence and workforce-planning models are the harder and more consequential cases. On their face, they sit outside the natural-persons framing of Annex III point 4. They operate on public and aggregate signals, and they report on skills, markets, roles, and scenarios rather than on candidates or workers. Where that description holds in full, a defensible position is that such a system is either outside the employment category altogether or, if an Annex III nexus is arguable, within the Article 6(3) derogation as a preparatory or advisory task that does not materially influence any individual decision. On this basis, a workforce-intelligence system can be characterised as governed strategic decision support rather than high-risk employment AI.

That position is defensible, but it is not automatic. Three tensions must be confronted honestly. The first is the 'materially influencing the outcome of decision-making' test in Article 6(3). Even guidance pitched at the level of a role can shape individual outcomes downstream: a recommendation to reduce a function eventually resolves into decisions about the people in it. The derogation speaks to the system's own function rather than to every remote consequence of acting on its advice, but a system whose outputs are

tightly coupled to individual actions invites a less comfortable reading. The second tension is the profiling proviso. Any feature that infers characteristics of identifiable natural persons, whether by de-anonymising an aggregate, scoring named individuals in a competitor's organisation, or resolving a 'role' to a single incumbent, engages the always-high-risk rule and collapses the derogation. The third is that classification under the AI Act does not displace other law. The profiling and automated-decision provisions of the GDPR, and fundamental-rights considerations more broadly, apply independently and may bite even where the AI Act does not.

The favourable classification of Archetypes C and D is, therefore, conditional on how they are built and used. The same broad idea of sensing the external labour market to inform strategy can be implemented as a clean aggregate advisory system, or as something that quietly profiles individuals and drives their treatment. The Act's criteria reward the former and penalise the latter. That contrast motivates a shift from arguing about data type to specifying architecture, which is the subject of the next section.

5. Five architectural properties that shift classification and defensibility

The following five properties are design choices that, taken together, make a workforce-intelligence system's regulatory position clearer and more defensible. They do not exempt a system by fiat; no design decision can override the Act's criteria. Each property does three things. It reduces the likelihood that a system falls within the Annex III employment scope. It strengthens an Article 6(3) non-high-risk assessment and the documentation required to support it under Article 6(4). Should the system be high-risk after all, it reduces residual risk against the high-risk obligations. The properties are stated as requirements a designer can implement and an auditor can verify.

5.1. The aggregation boundary

The system's outputs concern roles, functions, populations, and markets, and never identify, rank, or profile a natural person, whether directly or by trivial re-identification. This is the single most determinative property, because it keeps the system clear of both the natural-persons framing of Annex III point 4 and the always-high-risk profiling proviso in Article 6(3). Implementation means enforcing minimum aggregation thresholds, prohibiting outputs that resolve to a single incumbent, and testing outputs for re-identification risk as a standing control rather than a one-off. Where the aggregation

boundary is architecturally enforced, so that the system cannot emit individual-level inferences even if asked, the strongest classification argument is available.

5.2. Determinism and reproducibility

The discovery and normalisation of signals is deterministic and version-pinned, and sits ahead of any probabilistic or large-language-model layer, so that a given input and configuration reproduce a given output. This property speaks directly to the accuracy and robustness expectations of Article 15 and to the record-keeping logic of Article 12. It is also what makes an Article 6(4) assessment credible: a classification judgement can be tied to a specific, reproducible system state rather than to a moving target. Implementation means pinning pipeline and model versions, recording configurations, and being able to replay a prior run and obtain the same result. Placing deterministic processing before generative components confines non-determinism to a bounded, explainable part of the system.

5.3. Explain-only, read-only outputs

The system advises and explains; it does not act. Its outputs are read-only with respect to operational HR systems, it surfaces the reasoning and a calibrated confidence indication behind each output, and it has no write-back path that executes an employment decision. This property most directly supports the decision-support characterisation under Article 6(3). A system that improves or prepares a human assessment, and that is not meant to replace or influence that assessment without proper human review, aligns with the derogation's conditions. It also supports the human-oversight requirement of Article 14, by ensuring that a human always stands between the system's output and any consequence for a person.

5.4. Human-in-the-loop escalation

A named human owner is accountable for the system's outputs, with the ability to override them, defined escalation thresholds, and a documented review cadence. This property operationalises Article 14's human-oversight requirement and Article 26's deployer obligations. It converts 'a human decides' from an assurance into an auditable practice. Implementation means assigning ownership explicitly, recording overrides and their rationale, and reviewing the system's outputs and their use on a defined schedule rather than only when something goes wrong.

5.5. Provenance and end-to-end logging

Sources, transformations, and outputs are logged and auditable across the whole pipeline, so that any output can be traced back to the signals and steps that produced it. This property answers the record-keeping duty of Article 12, the automatic-logging expectation of Article 19, the data-governance and source-quality expectations of Article 10, and the technical-documentation requirement of Article 11. It also underwrites the other four properties: an aggregation boundary, a deterministic pipeline, and an explain-only interface are only as trustworthy as the audit trail that demonstrates they held in a specific case. Provenance is what turns the informal slogan 'logged, deterministic, reproducible' into evidence.

These properties are mutually reinforcing. The aggregation boundary keeps the system clear of the individual-level triggers. Determinism and provenance make that boundary demonstrable. The explain-only interface and human escalation ensure that even a well-founded output cannot become an unreviewed decision about a person. A system that exhibits all five is not only more likely to be classified as decision support; if classified as high-risk, it is also closer to meeting the high-risk obligations than a system designed without them.

The five properties are necessary but not sufficient. A system that meets all of them can still cause fundamental-rights harm through group-level advice that stigmatises a population, or through downstream operationalisation that couples aggregate guidance tightly to individual actions. The properties therefore presume a use-governance discipline around the system: assurance at procurement, effectiveness and bias testing across the lifecycle, monitoring after deployment, and named human roles that own the interpretation and consequences of the system's outputs. That discipline is the subject of Section 7.

6. A reference architecture and compliance checklist

The five properties can be embodied in a layered reference architecture. The description below is deliberately generic, so that it characterises a design pattern rather than any particular product.

Layer 1, external signal ingestion. Public and aggregate labour-market signals are collected with their provenance recorded and their licensing and permissions respected. Only aggregate or public sources are admitted; internal employee records are out of scope by design, which keeps the system clear of the employee-level data that draws Archetypes A and B toward the high-risk tier.

Layer 2, deterministic forensic discovery and normalisation. Signals are resolved, de-duplicated, and normalised through a deterministic, version-pinned process before any probabilistic component sees them. This layer establishes reproducibility and produces the lineage records on which later auditability depends.

Layer 3, aggregate intelligence and modelling. Skills, roles, markets, and scenarios are modelled at population level. Where probabilistic or language-model components are used, they operate on aggregate constructs, and aggregation thresholds are enforced so that outputs cannot resolve to an individual.

Layer 4, explain-only decision-support interface. Outputs are presented read-only, with reasoning and confidence surfaced, and with no path that writes back into systems that execute employment actions. A human owner interprets, overrides, and decides.

Layer 5, governance and audit spine. Cutting across the other layers, this spine logs sources, transformations, and outputs; records classification assessments and overrides; and holds the technical documentation. It should be built first, because it is what makes every other claim demonstrable.

The same properties can be expressed as an operational checklist that a design, procurement, or governance function can apply to a candidate system. Table 2 states each control as a question, names the provision of the Act it speaks to, and identifies the evidence a team should retain. The checklist is a working instrument rather than a definitive compliance determination. Several items, particularly those bearing on profiling and on material influence over individual decisions, require judgement that should be exercised with qualified legal input.

Control question	Relevant AI Act anchor	Evidence to retain
Do any outputs identify, rank, or profile a natural person, directly or by trivial re-identification?	Annex III(4); Art. 6(3) profiling carve-out	Data-flow map; re-identification test results; output schema showing role / population level only.
Are inputs public or aggregate labour-market signals rather than internal employee records?	Annex III(4) scope; Art. 10 (data governance)	Source inventory with provenance; licensing / permission records; data-minimisation notes.
Is discovery and normalisation deterministic and version-pinned before any probabilistic or LLM layer?	Art. 12 (record-keeping); Art. 15 (accuracy, robustness)	Pipeline version manifest; replay logs; fixed-seed reproduction of a prior run.
Does the system only explain and advise, with no write-back that executes an employment action?	Art. 6(3)(a), (c); Art. 14 (human oversight)	Interface spec showing read-only outputs; absence of decision-execution integrations.
Is a named human owner accountable, with override and a defined review cadence?	Art. 14 (human oversight); Art. 26 (deployer duties)	RACI; escalation thresholds; review-meeting records; override log.
Are sources, transformations, and outputs logged and auditable end to end?	Art. 12; Art. 19 (automatic logs); Art. 11 (technical documentation)	Immutable audit trail; lineage records; technical documentation pack.
If self-classified as not high-risk, is the Art. 6(3) assessment documented and registration prepared?	Art. 6(4); Art. 49(2) (registration)	Written non-high-risk assessment; registration record; sign-off by governance owner.
Has a fundamental-rights and GDPR overlay been assessed independently of AI Act tiering?	Art. 27 (FRIA, where applicable); GDPR Arts. 22, 35	FRIA or rationale for non-applicability; DPIA; profiling assessment under GDPR.

Table 2. Operational classification and governance checklist for workforce-intelligence systems.

A short worked illustration shows the checklist in use. Consider a system that maps task-level

automation exposure across research-and-development roles in a sector and produces ranked, explainable workforce actions at the level of role families. Designed to the five properties, it draws only on public role and task taxonomies and public labour-market signals. It resolves and normalises them deterministically. It models exposure at the level of role families and enforces aggregation thresholds. It presents ranked actions read-only, with reasoning and confidence, to a human planner who decides. It logs the whole chain. On the checklist, such a system reports no individual-level outputs, uses aggregate public inputs, processes them reproducibly, writes back nothing, is owned and reviewed by a named human, and is auditable end to end. That configuration supports a documented Article 6(3) assessment that the system is decision support rather than high-risk employment AI. A variant that scored named individuals, or wrote recommendations directly into a system that actioned them, would not.

7. From framework to practice: assurance, monitoring, and the operating model

The reference architecture and checklist of Section 6 describe a system in the abstract. Deploying such a system inside an organisation, whether built in-house or procured from a vendor, requires operational practices that make the five architectural properties verifiable, keep them intact after go-live, and place named humans in the roles the Act ascribes to providers and deployers. This section sets out those practices in five parts: how to assure a vendor product against the framework (7.1); how to test effectiveness across the lifecycle (7.2); how to test for bias and fairness (7.3); how to monitor and respond after deployment (7.4); and how to staff and consult around the whole regime (7.5).

7.1. Vendor and product assurance: a procurement due-diligence protocol

Most workforce-intelligence capability inside large organisations is not built from scratch. It is delivered by categories of enterprise product: core human-capital management suites and their embedded AI features for recruiting, matching, and internal mobility; standalone talent-intelligence platforms that sense external labour-market signals; AI-enabled applicant-tracking systems; skills-inference engines; and workforce-planning or AI-exposure modelling tools. Procurement is therefore where classification and defensibility are, in practice, either secured or lost. A vendor label of 'decision support' does not settle the question, because the classification depends on the architecture and use of the product, not on the label placed on it.

Two gates make procurement disciplined. The first is a classification gate. Before commercial evaluation begins, the candidate system is mapped to the archetype taxonomy (Table 1) and assessed against the five architectural properties (Section 5). The output of this gate is a written pre-purchase classification memo naming the archetype, the applicable Article 6(3) reasoning, and any residual profiling risk under the always-high-risk proviso. The memo is signed by an accountable owner and retained as evidence.

The second is an assurance gate. Vendor claims are matched against evidence in three artefacts that the algorithmic-auditing literature has established as minimum documentation for AI systems: a model card describing intended use, training data, evaluation metrics, and known limitations[7]; a datasheet describing the training and evaluation data sources, provenance, and known distributional properties[8]; and an internal or independent audit report covering effectiveness, bias, and robustness[9]. Where the system is a high-risk system under Annex III, the vendor supplies the technical documentation required by Article 11 and the deployer instructions required by Article 13. Where the vendor asserts non-high-risk status, the Article 6(4) assessment must accompany that assertion.

Practical due-diligence questions at the assurance gate include: which archetype does the product realise; what individual-level outputs, if any, can the product emit and under what configuration; where are the aggregation thresholds implemented and can they be inspected; is the discovery pipeline deterministic and version-pinned; what training and evaluation datasets underlie any model, and how have they been audited for representational bias; what are the vendor's post-market monitoring and incident-reporting arrangements under Articles 72 and 73; and what happens to logs and models if the contract terminates. A vendor unable to answer these questions on paper is not ready to be procured for an Annex III use case.

7.2. Effectiveness testing across the lifecycle

Article 15 requires that high-risk AI systems achieve appropriate levels of accuracy, robustness, and cybersecurity, and that those levels are declared in the accompanying instructions. Even where a workforce-intelligence system is classified as decision support under Article 6(3), effectiveness testing remains the empirical basis for the reliability the derogation presumes. The relevant metrics are established in the machine-learning and industrial-organisational psychology literatures, and the operating principles are consolidated in the NIST AI Risk Management Framework[10] and the AI-management-system standard ISO/IEC 42001:2023[11].

Predictive validity measures the correspondence between the system's output and the outcome it purports to inform, whether that outcome is a hiring decision, a promotion, a skills match, or a future

labour-market state. For classification outputs, discrimination is captured by the area under the receiver-operating-characteristic curve (AUROC) or, where the positive class is rare, by the area under the precision-recall curve (AUPRC). Ranking outputs are additionally evaluated with normalised discounted cumulative gain (nDCG) or mean average precision (MAP).

Calibration measures whether stated confidences match empirical frequencies. Two metrics dominate. The Brier score is defined, for a binary outcome $y \in \{0, 1\}$ and predicted probability $\hat{p}$, as $BS = (1/n) \Sigma_i (\hat{p}_i - y_i)^2$, with lower values indicating better calibration. The expected calibration error (ECE) partitions predictions into B bins by predicted probability and computes $ECE = \Sigma\_b (n\_b / n) \cdot |acc(b) - conf(b)|$, where acc(b) and conf(b) are the empirical accuracy and mean confidence in bin b.

Robustness measures how much output quality degrades under distribution shift, input perturbation, and adversarial noise. It is evaluated by out-of-time and out-of-distribution tests, by controlled perturbation of inputs, and by red-teaming exercises. Reliability is measured by test-retest stability across time and by human-machine agreement studies that compare system outputs with expert judgements on held-out cases. The overall regime should follow the sociotechnical framing of Selbst et al.[12]: effectiveness is a property of the system as used, not of the model in isolation.

These metrics are declared and monitored at three lifecycle points: model qualification before release; re-qualification after any change to data, model, or configuration; and continuously in production against pre-defined drift and accuracy thresholds that trigger retraining or rollback. For workforce-intelligence systems designed to Section 5′s properties, the deterministic-pipeline property makes model qualification reproducible, and the provenance property makes continuous monitoring an auditable trail rather than a scattering of dashboards.

7.3. Bias and fairness testing

Bias testing addresses whether a system distributes outputs equitably across protected and other sensitive groups. It is required implicitly by Article 10 on data governance, which demands examination of possible biases likely to affect health, safety, or fundamental rights. It is required explicitly by adjacent regimes that any global workforce-intelligence deployment will overlap: New York City Local Law 144, which mandates independent bias audits of automated employment decision tools[13]; the U.S. Equal Employment Opportunity Commission′s 2022 guidance on the use of algorithmic tools under the Americans with Disabilities Act[14]; ISO/IEC TR 24027:2021 on bias in AI systems[15]; and the NIST AI Risk

Management Framework[10]. Legal scholarship on hiring algorithms has been consistent in identifying bias detection as both technically necessary and legally salient[16][17].

For a system that classifies or scores individuals (Archetype A, and Archetype B where it profiles), four families of fairness metric are standard and should be reported together. Let $A \in \{0, 1\}$ denote a sensitive attribute, $\hat{Y} \in \{0, 1\}$ the system's decision, and $Y \in \{0, 1\}$ the ground-truth outcome.

- Demographic parity, also called statistical parity: $P(\hat{Y}=1 \mid A=1) \approx P(\hat{Y}=1 \mid A=0)$. Its practical form in employment law is the adverse impact ratio, $AIR = P(\hat{Y}=1 \mid A=1) \div P(\hat{Y}=1 \mid A=0)$, evaluated against the four-fifths rule ($AIR \geq 0.8$).
- Equal opportunity[18]: $P(\hat{Y}=1 \mid Y=1, A=1) \approx P(\hat{Y}=1 \mid Y=1, A=0)$. True positive rates are equal across groups.
- Equalised odds: equal true positive and false positive rates across groups, $P(\hat{Y}=1 \mid Y=y, A=1) \approx P(\hat{Y}=1 \mid Y=y, A=0)$ for $y \in \{0, 1\}$.
- Calibration parity: $P(Y=1 \mid \hat{p}, A=1) \approx P(Y=1 \mid \hat{p}, A=0)$ for all confidence scores $\hat{p}$. Predicted probabilities carry the same empirical meaning across groups.

These metrics are mutually incompatible in general: except in degenerate cases, no classifier can satisfy demographic parity, equalised odds, and calibration simultaneously[19][20]. Bias testing therefore begins with a documented choice of which metric to prioritise, justified against the specific decision the system informs, together with the trade-offs the choice implies. Intersectional testing across combinations of sensitive attributes is standard practice and often reveals disparities invisible at the marginal level.

For workforce-intelligence systems designed to the aggregation boundary (Archetypes C and D), individual-level fairness metrics are not applicable because there are no individual outputs to test. Bias risk instead migrates to two upstream locations. Representational bias in the input labour-market signals concerns the over- or under-representation of geographies, industries, seniority levels, and demographic proxies encoded in public postings; a concrete and well-documented form is self-selection bias in which posting platforms, geographies, and functions are unevenly present in the pool from which signals are drawn. Modelling bias concerns how skills, roles, and scenarios are constructed and weighted. Testing at these upstream locations uses distributional coverage metrics (population coverage, source diversity, and known-gap analyses), together with counterfactual probes that ask whether outputs change materially under plausible re-weightings of the input population. Selbst et al.[12] is the standard reference for the point that fairness is a property of a sociotechnical system, not of a model, and both individual-facing and aggregate systems inherit that framing.

Bias testing is conducted at three points: pre-deployment on a representative evaluation set; at re-qualification after any model or data change; and periodically in production against a fixed evaluation protocol. Results are recorded in the model card and audit report supplied to deployers and, where applicable, to authorities.

7.4. Post-deployment monitoring, incident reporting, and re-classification

Classification is not a one-time event. Article 72 requires providers of high-risk AI systems to establish and document a post-market monitoring system that actively and systematically collects, documents, and analyses relevant data on the system's performance throughout its lifecycle. Article 73 requires reporting of serious incidents to national authorities. Both provisions apply directly to high-risk systems, and for systems self-classified as non-high-risk under Article 6(3) they motivate an equivalent voluntary regime, because the Article 6(4) assessment is defensible only so long as its factual premises continue to hold.

A monitoring regime for a workforce-intelligence system covers three axes. Performance monitoring tracks the effectiveness metrics defined in Section 7.2 against thresholds that trigger investigation, retraining, or rollback. Drift monitoring detects data drift in the input signals (changes in the distribution of postings, sources, or geographies), concept drift in the mapping from signals to outcomes (skills taxonomies shifting under labour-market change), and label drift where applicable. Bias monitoring re-runs the fairness testing of Section 7.3 on production data at a defined cadence. The technical patterns are well established in the ML-operations literature, whose canonical warning about the accumulated cost of unmonitored systems is Sculley et al.[21] on hidden technical debt.

Two triggers matter specifically for classification. The first is scope creep: a product extension that begins to identify individuals, resolve a role to an incumbent, or write back into a system that executes an action collapses the Article 6(3) argument and requires re-classification as high-risk. The second is a material shift in the input distribution such that outputs no longer represent the same underlying phenomenon; this triggers re-qualification rather than re-classification, but has similar governance consequences. Both triggers are listed in the technical documentation with named owners and escalation paths.

Serious incidents in this domain include a discovered profiling capability that was not disclosed at classification, a demonstrable adverse impact on identifiable persons downstream of aggregate advice, or a security incident affecting model, training data, or logs. The reporting timelines and content are

specified by Article 73 for high-risk systems; internal parallel procedures for non-high-risk systems reduce the risk that a late discovery becomes a late disclosure.

7.5. The AI governance operating model and worker-representative consultation

The architectural properties do not implement themselves, and the assurance, testing, and monitoring practices above depend on named humans in defined roles. Consistent with the structural, relational, and procedural framing of responsible AI governance[2], a workforce-intelligence deployment involves a small number of functions with clearly separated responsibilities.

A senior AI governance lead, sometimes titled AI ethics officer, owns the classification decision, the Article 6(4) assessment where applicable, and the enterprise register of AI systems. The Data Protection Officer, already required under the GDPR, owns the profiling and automated-decision assessment under GDPR Articles 22 and 35 and coordinates the fundamental-rights impact assessment under Article 27 of the Act where applicable. Legal counsel and compliance own the mapping from the Act's articles to internal policy and to contractual terms with vendors. Information security owns the cybersecurity portion of Article 15 and the integrity of the logs described in Article 12.

On the technical side, ML and AI engineers own the effectiveness and bias-testing pipelines and the deterministic-processing property. Forward-deployed engineers translate business and workforce-planning questions into system configurations, run reproducible analyses used by decision-makers, and act as the SME bridge between technical implementation and business use. HR business partners and workforce-planning subject-matter experts act as decision owners under Article 26, interpreting outputs, exercising overrides, and documenting the reasoning behind their decisions. Internal audit periodically re-runs the checklist of Section 6 against evidence.

Two roles are frequently underplayed. The first is the worker-representative body, whether an internal works council, a European Works Council, or an equivalent employee representation forum. Article 26(7) requires deployers who are employers to inform worker representatives and affected workers before a high-risk AI system is put into service. Both this obligation and the underlying labour-law principles are engaged even where a system is self-classified as non-high-risk. Consultation is best treated as a deployment gate rather than an after-the-fact notification, because it surfaces classification questions that a purely technical review will miss.

The second underplayed role is a standing governance forum that convenes the functions listed above at defined intervals, reviews the register of AI systems, and signs off on new classifications and material

changes. Without such a forum, classification memos and monitoring reports accumulate without adjudication, and re-classification triggers are less likely to be acted on in time. A compact responsibility mapping is provided in Table 3.

Activity	Responsible	Accountable	Consulted	Informed
Pre-purchase classification memo	AI governance lead	Business owner	Legal, DPO, ML engineering	HR leadership
Article 6(4) non-high-risk assessment	AI governance lead	Business owner	Legal, DPO	Executive committee
Effectiveness testing (design and production)	ML / AI engineering	AI governance lead	Workforce-planning SME, forward-deployed engineer	Business owner
Bias and fairness testing	ML / AI engineering	AI governance lead	DPO, legal, HR business partner	Worker representatives (results)
Post-market monitoring	ML / AI engineering	AI governance lead	Information security, forward-deployed engineer	Business owner
Serious-incident reporting	Legal / AI governance lead	Business owner	DPO, information security	Executive committee, authority
Worker-representative consultation	HR business partner	Business owner	AI governance lead, legal	ML / AI engineering
Internal audit of the checklist	Internal audit	Audit committee	AI governance lead	Executive committee

Table 3. Indicative RACI for workforce-intelligence deployment

The RACI is indicative rather than prescriptive; scale and structure vary with organisation size and geography. The invariant is separation of the classification, testing, decision, and audit roles, so that no single function combines the design of a system with the assessment of its regulatory position and the review of its operation.

8. Implications and a research agenda

The research question posed in Section 1 asked how a workforce-intelligence system should be built, procured, tested, monitored, and governed so that its classification is clear and defensible at deployment and remains so through operation. The taxonomy, five properties, reference architecture, checklist, and practice layer are the paper's answer. This section draws out the implications of that answer for three audiences and then sets out a research agenda.

8.1. For HR and workforce-planning leaders

The first implication is that classification is a design and governance decision, not merely a procurement checkbox. Whether a workforce-intelligence capability is high-risk depends on how it is built and used, and leaders who treat the question as settled by a vendor's label expose themselves to a contestable position. The second is documentation: where a system is judged not to be high-risk under Article 6(3), that judgement must be written down before deployment and the registration obligation observed. The third is discipline against scope creep. The favourable classification of external workforce intelligence depends on the aggregation boundary holding; a well-meaning extension that begins to profile named individuals can move a system across the high-risk line without anyone deciding to do so. The procurement due-diligence protocol of Section 7.1 and the operating-model roles of Section 7.5 make this discipline operational.

8.2. For system designers and vendors

For those who build these systems, the five properties are design requirements rather than after-the-fact controls. The governance and audit spine should be built first, because it is what makes every other property demonstrable. Designing the aggregation boundary as an architectural constraint, so that the system cannot emit individual-level inferences, is more defensible than relying on policy alone. Placing deterministic processing ahead of generative components confines non-determinism to a bounded, explainable region and makes reproducibility a property of the system rather than an aspiration. The effectiveness and bias-testing regimes of Sections 7.2 and 7.3, together with the monitoring cadence of Section 7.4, translate the design into an evidence stream that supports classification and ongoing quality assurance in parallel.

8.3. For governance, risk, and legal functions

Governance functions should treat the profiling proviso as a bright line in their review. They should assess fundamental-rights and data-protection exposure independently of AI Act tiering, since the GDPR's profiling and automated-decision provisions apply on their own terms. Where a fundamental-rights impact assessment is required, the provenance and logging property supplies much of the evidentiary base it needs. Coordination across the AI Act, the GDPR, and internal governance is more effective when the system is designed to produce audit evidence as a by-product of operation rather than on demand. The indicative RACI in Table 3 makes accountability concrete, and Article 26(7) consultation with worker representatives is treated as a design gate rather than an after-the-fact notification. Governance review should also consider candidates and workers whose data flows into the aggregate signal pool as stakeholders in their own right: even where a system does not decide about them, their information underwrites its outputs, and their expectations under GDPR transparency and labour-law consultation carry independently of AI Act tiering.

8.4. Limitations

This paper is analytical and design-oriented, not empirical or doctrinal. Its classification arguments are offered as reasoning to be tested, not as legal conclusions, and they precede the Commission's anticipated guidance and practical examples, which may refine or unsettle particular positions. The reference architecture and the practice layer reflect a designer's vantage rather than a survey of deployed systems. The analysis concentrates on the employment category rather than on the biometric and other Annex III categories that a workforce system must also avoid triggering. The fairness metrics and monitoring practices summarised in Section 7 are curated rather than exhaustive, and their application to any given system requires domain-specific choices. These limitations define the empirical work that should follow.

8.5. A research agenda

Several lines of work would strengthen the account. A structured evaluation of the framework's artefacts, following the Framework for Evaluation in Design Science (FEDS) proposed by Venable et al.[5], would establish which of the five properties and which of the practice-layer controls are essential and which are contingent. Expert validation, through structured interviews or a Delphi study with workforce-planning leaders, system designers, and AI-governance and legal specialists, would test the archetype taxonomy and the properties against practice. A cross-jurisdiction comparison would situate the framework against

other regimes and voluntary standards, clarifying which properties are regime-specific and which are general. A focused study of the determinism-and-auditability property would examine its costs and its trade-offs against model performance. Case studies of organisations that have implemented workforce-intelligence systems under the Act, and longitudinal work on how classification judgements fare under scrutiny, would convert the framework from a design proposal into an evidence-based standard. This agenda also marks the path from the present working paper to a peer-reviewed treatment.

9. Conclusion

The EU AI Act makes employment AI high-risk. It does so, however, through language framed around identifiable natural persons and decisions about them, and it leaves room, through Article 6(3), for systems that advise rather than decide. External workforce-intelligence systems, which read aggregate labour-market signals to inform strategy at the level of roles and populations, can be built to sit on the decision-support side of that line, but only if they are built deliberately and sustained deliberately after deployment. This paper has argued that classification turns on architecture rather than on data type. It has offered a taxonomy of four archetypes, a classification analysis anchored in the Act's own criteria, and five architectural properties: an aggregation boundary, deterministic and reproducible processing, explain-only outputs, human-in-the-loop escalation, and end-to-end provenance logging. Expressed as a reference architecture and a working checklist, and extended into a practice layer for vendor assurance, effectiveness and bias testing, post-deployment monitoring, and the operating-model roles that make the framework operational, these give the people who design, procure, and govern such systems a practical instrument for the Act's phased enforcement, and a foundation for the empirical work that responsible workforce intelligence now needs.

Notes

This working paper offers scholarly analysis and design guidance. It is not legal advice, and its classification arguments should be reviewed by qualified counsel before being relied upon for compliance decisions.

JEL Classification: K23 (Regulated Industries and Administrative Law); K31 (Labor Law); J24 (Human Capital; Skills; Occupational Choice; Labor Productivity); M15 (IT Management); M54 (Labor Management); O33 (Technological Change: Choices and Consequences; Diffusion Processes).

Statements and Declarations

Funding

This research received no external funding. The manuscript was developed independently by the author using publicly available literature, frameworks, and regulatory guidance.

Potential Competing Interests

The author is employed by Novartis Healthcare Pvt. Ltd. The views and opinions expressed in this article are solely those of the author and do not necessarily represent the views, policies, or positions of Novartis. No commercial product, service, or vendor is evaluated or promoted in this manuscript.

References

1. ^Mökander J, Sheth M, Watson DS, Floridi L (2023). "The Switch, the Ladder, and the Matrix: Models for Classifying AI Systems." Minds Mach. **33**(1):221–248. doi:10.1007/s11023-022-09620-y.
2. a, bPapagiannidis E, Mikalef P, Conboy K (2025). "Responsible Artificial Intelligence Governance: A Review and Research Framework." J Strategic Inf Syst. **34**(2):101885. doi:10.1016/j.jsis.2024.101885.
3. ^Hevner AR, March ST, Park J, Ram S (2004). "Design Science in Information Systems Research." MIS Q. **28**(1):75–105. doi:10.2307/25148625.
4. a, bPeffers K, Tuunanen T, Rothenberger MA, Chatterjee S (2007). "A Design Science Research Methodology for Information Systems Research." J Manage Inf Syst. **24**(3):45–77. doi:10.2753/MIS0742-1222240302.
5. a, bVenable J, Pries-Heje J, Baskerville R (2016). "FEDS: A Framework for Evaluation in Design Science Research." Eur J Inf Syst. **25**(1):77–89. doi:10.1057/ejis.2014.36.
6. ^van den Heuvel S, Bondarouk T (2017). "The Rise (and Fall?) of HR Analytics: A Study into the Future Application, Value, Structure, and System Support." J Organ Eff People Perform. **4**(2):157–178. doi:10.1108/JOEPP-03-2017-0022.
7. ^Mitchell M, Wu S, Zaldivar A, Barnes P, Vasserman L, Hutchinson B, Spitzer E, Raji ID, Gebru T (2019). "Model Cards for Model Reporting." In Proceedings of the Conference on Fairness, Accountability, and Transparency (FAccT '19), 220–229. doi:10.1145/3287560.3287596.
8. ^Gebru T, Morgenstern J, Vecchione B, Vaughan JW, Wallach H, Daumé III H, Crawford K (2021). "Datasheets for Datasets." Commun ACM. **64**(12):86–92. doi:10.1145/3458723.

9. ^Raji ID, Smart A, White RN, Mitchell M, Gebru T, Hutchinson B, Smith-Loud J, Theron D, Barnes P (2020). "Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing." In Proceedings of the Conference on Fairness, Accountability, and Transparency (FAccT '20), 33–44. doi:10.1145/3351095.3372873.
10. a, bNational Institute of Standards and Technology (2023). "Artificial Intelligence Risk Management Framework (AI RMF 1.0)." NIST AI 100-1. doi:10.6028/NIST.AI.100-1.
11. ^International Organization for Standardization / International Electrotechnical Commission (2023). "ISO/IEC 42001:2023 Information Technology, Artificial Intelligence, Management System." ISO.
12. a, bSelbst AD, Boyd D, Friedler SA, Venkatasubramanian S, Vertesi J (2019). "Fairness and Abstraction in Sociotechnical Systems." In Proceedings of the Conference on Fairness, Accountability, and Transparency (FAccT '19), 59–68. doi:10.1145/3287560.3287598.
13. ^New York City (2023). "Local Law 144 of 2021: Automated Employment Decision Tools." New York City Administrative Code, Title 20, Chapter 5, Subchapter 25.
14. ^U.S. Equal Employment Opportunity Commission (2022). "The Americans with Disabilities Act and the Use of Software, Algorithms, and Artificial Intelligence to Assess Job Applicants and Employees." Technical Assistance Document EEOC-NVTA-2022-2.
15. ^International Organization for Standardization / International Electrotechnical Commission (2021). "ISO/IEC TR 24027:2021 Information Technology, Artificial Intelligence (AI), Bias in AI Systems and AI Aided Decision Making." ISO.
16. ^Barocas S, Selbst AD (2016). "Big Data's Disparate Impact." Calif Law Rev. **104**(3):671–732. doi:10.15779/Z38BG31.
17. ^Ajunwa I (2020). "The Paradox of Automation as Anti-Bias Intervention." Cardozo Law Rev. **41**(5):1671–1742.
18. ^Hardt M, Price E, Srebro N (2016). "Equality of Opportunity in Supervised Learning." In Advances in Neural Information Processing Systems (NeurIPS 2016), 29, 3315–3323.
19. ^Kleinberg J, Mullainathan S, Raghavan M (2017). "Inherent Trade-offs in the Fair Determination of Risk Scores." In 8th Innovations in Theoretical Computer Science Conference (ITCS 2017), Vol. 67, 43:1–43:23. doi:10.4230/LIPIcs.ITCS.2017.43.
20. ^Chouldechova A (2017). "Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments." Big Data. **5**(2):153–163. doi:10.1089/big.2016.0047.

21. ^*Sculley D, Holt G, Golovin D, Davydov E, Phillips T, Ebner D, Chaudhary V, Young M, Crespo J-F, Dennison D (2015). "Hidden Technical Debt in Machine Learning Systems." In Advances in Neural Information Processing Systems (NeurIPS 2015), 28, 2503–2511.*

Declarations

Funding: This research received no external funding. The manuscript was developed independently by the author using publicly available literature, frameworks, and regulatory guidance.

Potential competing interests: The author is employed by Novartis Healthcare Pvt. Ltd. The views and opinions expressed in this article are solely those of the author and do not necessarily represent the views, policies, or positions of Novartis. No commercial product, service, or vendor is evaluated or promoted in this manuscript.