AIQA Global

The Standard for Enterprise AI Quality

aiqaglobal.com
info@aiqaglobal.com

THE MEASURED VIEW

AIQA Global's series on the ideas, standards, and market forces shaping enterprise AI governance. In a field defined by speed and speculation, these articles offer something different: informed perspective grounded in the discipline of measurement. Because the organizations that will lead in AI are the ones that can prove it.

When Everyone's Responsible, No One Is

A patchwork of self-attesting AI vendors leaves no one accountable. An independent rating locates the responsibility.

Issue #07 — June 2026 — AIQA Global, LLC

The AI Procurement Officer

Every week brings another AI vendor and another claim. The hard part was never adopting AI. It is knowing whose AI to trust — and who answers when the stitched-together whole fails.

A procurement officer sits with a stack of vendor decks and a decision to make. Forty of them, near enough. Every deck says the same thing in slightly different fonts: responsible AI, integrated, human-in-the-loop, fully compliant, "AI-powered" everything. Every vendor is the safer, faster, smarter choice. None are lying. And none can be ranked against the other on the strength of the page alone.

The careful buyer does more than read the decks. She sends a vendor questionnaire, scores the answers against an internal rubric, runs each tool through an AI due dilligence the organization is building. And, in the end, the decision is made based on which company describes itself more convincingly.

Buying Claims. Not Verified Quality.

The questionnaire is filled in by the vendor. The rubric grades the vendor's account of itself. The checklist standardizes the questions, not the evidence behind them. Diligence organizes self-description; it does not verify it and it cannot set one vendor's governance beside another's on a common scale. The instrument measures how completely a vendor answered – not how well it actually governs.

That is the problem in one image. Self-description is not evidence. And in AI procurement, self-description - however neatly it is scored - is most of what a buyer is handed. This is the ordinary, unglamorous shape of the AI governance problem in 2026. It does not announce itself as a crisis. It arrives as a purchasing decision — repeated dozens of times a year, made largely on the strength of each vendor's description of itself.

A description is easy to inflate - a problem that regulators have caught on to. Cue "AI Washing" – the term for inflated or false claims about how a company uses artificial intelligence - a practice that the Securities and Exchange Commission (SEC) now polices.

In March 2024 the agency fined two investment advisers for misrepresenting how they used AI, with civil penalties of $225,000 and $175,000. In January 2025 it charged a restaurant-technology company that had told investors its AI eliminated the need for human drive-thru order-taking - when the vast majority of those orders still ran through a person. A month later the agency stood up a dedicated Cyber and Emerging Technologies Unit to pursue exactly this kind of misconduct in AI.

The lesson generalizes. If a regulator with subpoena power will not take "we use AI" at face value, a procurement officer with forty decks and no enforcement authority certainly should not. The claim on the page is the beginning of due diligence, not the end of it.

The AIQ Quilt and The Problem of Many Hands

No organization buys a single AI. It builds a patchwork, stitched together across the value chain: a foundation model from one provider, a vendor's fine-tuned layer on top of it, a SaaS tool that orchestrates the workflow, a data broker that supplies inputs, an analytics add-on that interprets the outputs. Each supplier attests to its own square of the quilt — its model is safe, its layer is compliant, its tool is governed. No party holds an independent, comparable view of the whole. The seams between the squares, where most real failures happen, belong to no one.

This is not an edge case. Fragmentation is now the norm: enterprise buyers deliberately assemble multi-model, multi-vendor stacks, a pattern documented across the 2025 enterprise surveys from Andreessen Horowitz and Menlo Ventures, the latter estimating roughly $18 billion invested across the model, training, data, and orchestration layers in a single year. The stitched-together stack is the architecture buyers chose, not an accident they fell into.

There is an older idea underneath this, and naming it sharpens what is actually broken. Accountability theorists call it ***the problem of many hands***: the difficulty of assigning responsibility for a harm produced by many actors, none of whom did the whole thing. The philosopher Helen Nissenbaum brought the idea into computing in a 1996 essay, as the first of four barriers to accountability in computerized systems - the fourth being *ownership without liability*, the contractual move that disclaims responsibility for what you ship. A 2022 follow-up by A. F. Cooper and colleagues revisited those barriers for the machine-learning era, where the actors are more numerous and the causal chains harder to trace.

The point is not that AI created this problem. The point is that AI *industrializes* it. Every automated layer — every model, every fine-tuning pass, every orchestration tool, every data feed — is another hand. The principle is old. What is new is that the technology now manufactures the diffusion at scale, and builds it into the architecture rather than leaving it to circumstance.

The Failure No One Owns

Stitched-together systems fail in stitched-together ways. A hallucinated output, a mispriced risk, a wrongful denial — the failure rarely sits cleanly in one vendor's square. And when responsibility is shared across a chain, the predictable outcome is that no one holds it.

The deployer points to the vendor. The vendor points to the model provider. The model provider points to the training data. Every party can credibly gesture at another and each, in turn, is partly right. Diffuse accountability is not a failure of the chain. It is the default state of any chain assembled from self-attested parts.

Start with the cleanest example. In *Moffatt v. Air Canada* (2024 BCCRT 149), an airline's website chatbot gave a customer wrong information about bereavement fares. When the customer sought to be made whole, the airline argued — remarkably — that its own chatbot was "a separate legal entity" responsible for its own statements. British Columbia's Civil Resolution Tribunal refused the move and held the airline responsible for the information it published, by whatever mechanism it published it. The decision is a tribunal ruling rather than a court judgment, and the dollars at stake were small. But the principle it rejected is the one that matters: a deployer cannot disown the tool it chose to deploy.

The diffusion runs the other way too — from vendor to deployer. In *Louis v. SafeRent Solutions*, tenant applicants alleged that an algorithmic screening score discriminated against Black and Hispanic renters and housing-voucher holders. SafeRent's defense was that it could not be liable because it never made the final accept-or-deny decision; the landlord did. It "only scored." The court denied the motion to dismiss on those counts — the score still shaped who got access to housing, regardless of who clicked the final button — and the Department of Justice and the Department of Housing and Urban Development backed the plaintiffs. The case settled

in 2024 for roughly $2.28 million. Tellingly, the settlement requires that any future screening score SafeRent develops be validated by an independent third party before use — a private resolution arriving, on its own, at independent validation as the remedy.

The live frontier of this question is *Mobley v. Workday* (No. 3:23-cv-00770, N.D. Cal.), and it should be watched rather than summarized, because it is moving. The plaintiffs allege that an AI applicant-screening system produced age-discriminatory outcomes across the many employers that use it. Workday's position was that it is a software provider, not an employer, and therefore beyond the reach of employment-discrimination law. The court has so far allowed the claims to proceed on the theory that the vendor's AI could be an *agent* of the employers that delegated screening to it; it granted preliminary certification of an age-discrimination collective; and in a March 6, 2026 ruling it held that the relevant protections reach job applicants, not only current employees. Plaintiffs filed an amended complaint later that month, and the matter is now in discovery. Nothing here is decided — but the question itself is consequential, and the scale is the reason: the screening tools at issue touch a large share of the country's biggest employers, and the certified collective could reach into the millions. This case is best read as live and escalating.

One further data point belongs in the picture, with a distinction kept clean. In *EEOC v. iTutorGroup* (2023), an employer's recruiting software automatically rejected older applicants; the company paid $365,000 in what was the EEOC's first settlement of an AI-discrimination claim. The principle — automation is no excuse for a discriminatory outcome — is directly on point. But this was an employer configuring its *own* software, not a third party's product, so it is not another link in the vendor-to-deployer chain. It is the floor under the whole discussion: putting a decision through software does not launder it of responsibility.

Read together, the cases share a structure. Wherever the diffusion was allowed to stand, responsibility evaporated. Wherever a tribunal or court refused it, that body had to *manufacture* a fixed point — the deployer, the "agent" — after the harm had already occurred. The need to construct the fixed point after the fact is the proof that it was missing by design.

> *A self-attestation tells you what a vendor says about itself. An independent rating tells you what a third party found.*

It is worth being precise about how this differs from a point made earlier in this series. *AI Governance Is Not Optional* argued that organizations carry more AI exposure than they think. This is a different claim. The exposure is not merely larger than assumed — it has **no single owner by design**. The quilt was assembled so that every square has an attestor and the whole has none.

Indemnity is Not the Answer

The instinctive corporate response to chain risk is: contracts, SLAs, indemnities. Many AI vendors now offer one. Microsoft's Copilot Copyright Commitment, announced in 2023, and comparable commitments from Adobe and Google Cloud, are the prominent examples. They are real protections, and worth having.

They are also narrow. These indemnities largely cover **third-party intellectual-property and copyright claims** arising from a model's outputs — and even then only conditionally: the customer must use the vendor's built-in safety systems, must not knowingly infringe, and typically gets no coverage for trademark claims. What they do not cover is most of what actually goes wrong with deployed AI: hallucinated output relied upon as fact, mispriced risk, a wrongful denial of credit or housing or employment, the actions of an autonomous agent, a failure that emerges from the interaction of several vendors' systems, or a regulatory enforcement action.

The deeper limitation is structural, and it is the heart of the matter. An indemnity reallocates the **cost** of a failure — it decides who writes the check after the fact. It never locates the **responsibility** for it — who is accountable before anything goes wrong. Those are different questions, and only one of them is answered by a clause.

> *An indemnity reallocates the cost of a failure. It never locates the responsibility for it.*

Governance Locates Responsibility

Here is where the argument turns, and where it is easy to reach for the wrong solution. Confronted with a responsibility problem, organizations tend to add internal processes — another committee, another policy binding, another sign-off. But layering more processes onto a process problem does not produce accountability, it produces paperwork.

This is the moment when governance becomes a ‘dirty’ word. Treated as paperwork, it reads as bureaucracy and burden, a tax on speed. Treated correctly, it is the opposite. Governance, done well, is what makes accountability locatable — the difference between "we all did our best" and "here is who was responsible, and here is the evidence." It is what lets you answer "whose failure is this" with a name and a record rather than a shrug, before the question is asked in a deposition instead of a meeting.

What the situation requires is different in kind. An independent, comparable mark does the one thing that contracts and self-attestations structurally cannot: it places every party on a single axis and assigns each a fixed coordinate of responsibility. This is the distinction between a credit rating and a company's own marketing about its creditworthiness. A self-description is a first-party claim; a rating is a third-party finding. The difference is not tone or rigor — it is *who is speaking*, and whether they have a stake in the answer.

The mechanism is simple enough to state in a sentence. **Independence** — the assessor holds no interest in the system being rated —and **comparability** — every system is scored on one standardized scale. AIQA set out the conditions such a mark must satisfy in its Chicago Principles for Independent AI Assurance: that assurance be Independent, Measurable, Auditable, Comparable, Continuously Updated, and Accountable. Those six words describe the difference between a finding and a claim.

What a Fixed Point Changes

Restore the missing fixed point and the picture changes for each of the constituencies that depend on enterprise AI being trustworthy.

For **procurement**, the forty decks come with a selection signal. . The officer chooses on a third-party score rather than on the eloquence of a vendor's claims – a comparable floor under a decision that previously had none.

The same signal does further work downstream: during **integration**, it is a governance record, an outside account of who governs what, to what standard, held independently of any vendor's marketing; at the point of **failure**, it is an accountability anchor, the documented basis for locating responsibility rather than diffusing it.

For **boards**, AI risk acquires a locatable owner. Fiduciary oversight requires knowing who is answerable for a system before it fails; an independent assessment supplies that, in a form a board can actually govern against.

For **insurers**, the score is an accountability signal that underwriting can price. The recurring difficulty in pricing AI risk is the absence of standardized, comparable data; a third-party rating is the kind of input a carrier can build into a model.

For **regulators**, an independent mark gives an emerging legal allocation something to measure against. The EU AI Act already assigns responsibility to roles rather than leaving it to diffuse: Article 25 provides that a distributor, importer, deployer, or other party can be treated as a "provider" — assuming the provider obligations of Article 16 — in defined circumstances, such as putting its name on a high-risk system or substantially modifying it; Articles 16 through 20 set out the provider-obligation cluster, from quality-management systems to documentation, logging, and corrective action; and Article 26 sets the obligations of deployers. These high-risk obligations are phasing into force through 2026. The law is trying to answer the "who is the provider?" question. An independent, comparable assessment gives that answer something to be checked against.

This is the discipline AIQA Global built the AIQ™ Score to provide: an independent, quantitative rating of how well an organization governs its AI – expressed on a 0-200 scale across five weighted dimensions and 250 data points, assessed by a third party rather than asserted by the system under review. Not to slow the quilt's assembly — to make it bear weight.

Back in The Room

Return to the procurement officer and the stack of decks. The point was never to buy less AI, or to move more slowly than competitors. The officer who can rank the vendors on an independent score is not the cautious one in the building. She is the only one who can say, before anything goes wrong, where the responsibility for it sits.

The question facing boards and enterprise leaders is no longer whether to adopt AI, or how fast. It is quieter and harder: can you show that the AI you have assembled is well-governed — to a standard you did not set yourself?

A vendor's confidence is not evidence. An indemnity is not accountability. And a quilt is only as trustworthy as the proof that someone, independent of the people who sold you the squares, has checked the seams.

When responsibility is spread across the whole value chain, it does not get shared. It disappears. An independent mark is how you find it again.

See how the AIQ™ Score quantifies AI documentation, governance, and deployment: aiqaglobal.com/aiq-score

Sources

1. Helen Nissenbaum, "Accountability in a Computerized Society," *Science and Engineering Ethics* 2(1):25–42 (1996) — the canonical computing treatment of the "problem of many hands," named as one of four barriers to accountability.

2. A. F. Cooper et al., "Accountability in an Algorithmic Society: Relationality, Responsibility, and Robustness in Machine Learning," arXiv:2202.05338 (2022; also presented at FAccT) — revisits Nissenbaum's four barriers for the machine-learning era.
3. *Moffatt v. Air Canada*, 2024 BCCRT 149 (British Columbia Civil Resolution Tribunal, Feb. 14, 2024) — tribunal rejected the airline's "separate legal entity" characterization of its chatbot and held the deployer responsible.
4. *Louis et al. v. SafeRent Solutions*, No. 1:22-cv-10800 (D. Mass.) — motion to dismiss denied in part (July 2023); ~$2.275M settlement granted final approval Nov. 20, 2024; DOJ and HUD filed a statement of interest; settlement requires independent third-party validation of any future screening score.
5. *Mobley v. Workday, Inc.*, No. 3:23-cv-00770-RFL (N.D. Cal.) — second motion to dismiss denied (July 2024) on an "agent" theory; preliminary ADEA collective certification (May 16, 2025); ruling that the ADEA reaches job applicants (Mar. 6, 2026); amended complaint filed (Mar. 30, 2026); now in discovery. *Status current as of June 15, 2026 — re-verify before publication.*
6. *EEOC v. iTutorGroup, Inc.*, No. 1:22-cv-02565 (E.D.N.Y.) — $365,000 consent decree (2023); the EEOC's first settlement of an AI-discrimination claim involving an employer's own recruiting software.
7. U.S. Securities and Exchange Commission, *In re Delphia (USA) Inc.* and *In re Global Predictions, Inc.* (Mar. 18, 2024) — first AI-washing enforcement actions against investment advisers; combined $400,000 in civil penalties.
8. U.S. Securities and Exchange Commission, *In re Presto Automation Inc.* (Jan. 14, 2025) — settled charges that the company overstated its AI's ability to complete drive-thru orders without human intervention.
9. U.S. Securities and Exchange Commission — creation of the Cyber and Emerging Technologies Unit (Feb. 2025).
10. Microsoft, "Introducing the Microsoft Copilot Copyright Commitment" (Sept. 2023); comparable indemnities offered by Adobe and Google Cloud — coverage of third-party IP/copyright claims on outputs, subject to conditions; trademark generally excluded.
11. EU AI Act — Article 25 (responsibilities along the AI value chain), Articles 16–20 (provider obligations), Article 26 (deployer obligations). High-risk obligations phasing into force through 2026.
12. Andreessen Horowitz, "How 100 Enterprise CIOs Are Building and Buying Gen AI in 2025"; Menlo Ventures, "2025: The State of Generative AI in the Enterprise" — enterprise adoption of multi-model, multi-vendor stacks. *Confirm specific figures at source before publication.*
13. AIQA Global, "The Chicago Principles for Independent AI Assurance" (April 2026) — Independent, Measurable, Auditable, Comparable, Continuously Updated, Accountable.