The People-Vector Evidence Layer for AI Governance Frameworks

Mapping Behavioral Assessment to NIST AI RMF, ISO/IEC 42001, and the EU AI Act

PAICE.work PBC

June 2026 • Version 1.1

A reference for risk, compliance, and security officers in regulated industries

Executive Summary

AI governance frameworks have converged on a requirement that most organizations cannot yet satisfy: demonstrable evidence that the humans operating AI-enabled workflows exercise meaningful oversight. NIST AI RMF asks organizations to measure the human-AI configuration. ISO/IEC 42001 requires competent personnel and defined responsible-use controls within the AI management system. The EU AI Act requires actors to support AI literacy among the staff who operate AI, and makes effective human oversight a legal obligation for high-risk systems from December 2027. Each framework points at the same place: the behavior of the people in the loop. Almost no organization has an instrument that produces evidence there.

Compliance teams fill this gap with the artifacts they have: training-completion records, usage logs, and self-attestations. None of these is behavioral evidence. A completion certificate proves attendance, not competence. A usage log proves activity, not reliability. A self-attestation records what a person believes about their own practice, which is the least reliable signal of all. An auditor asking “show me that your people catch AI errors before those errors reach a client” cannot be answered with any of them.

PAICE (People + AI Collaboration Effectiveness) produces the missing artifact. It observes how a professional behaves when AI is uncertain, incomplete, or wrong, scores that behavior across five dimensions on a 0–1000 scale, and does so without retaining conversation content or participant identity. This paper maps PAICE output to the specific clauses of NIST AI RMF, ISO/IEC 42001, and the EU AI Act that concern the human vector, and shows how a compliance officer can place that output into an audit file. PAICE does not certify an organization against any framework. It supplies the People-vector evidence those frameworks ask for and that the rest of the governance stack does not generate.

This is the latest in PAICE’s whitepaper series and the first to address the AI-governance frameworks. Where the companion **Verifiable Human-AI Collaboration** paper (presented at NEARCON 2026) maps PAICE’s architecture to privacy and verifiability law (GDPR, CCPA/CPRA, the NIST Privacy Framework, with TEE-protected inference and on-chain attestation), this paper maps PAICE’s behavioral output to AI-governance obligations. The two axes are complementary, not duplicative; the Appendix lists the companion papers and which reviewer each is written for.

The Evidence Burden Has Shifted to Behavior

For most of the last decade, AI governance was a question about systems. Was the model documented? Was the data lineage recorded? Were access controls in place? The 2024–2026 wave of frameworks added a second axis that is harder to instrument: the conduct of the people

operating the system. The shift is explicit in the source texts. Human oversight is no longer assumed; it must be designed, evidenced, and in some jurisdictions proven to a regulator.

Three reference frameworks now anchor most regulated-industry programs. The **NIST AI Risk Management Framework** organizes practice into four functions (Govern, Map, Measure, and Manage), and its Measure function calls for methods and metrics that include the human-AI configuration and the trustworthiness characteristics of the deployed system. **ISO/IEC 42001:2023**, the first certifiable AI management system standard, requires competent personnel (Clause 7.2), defined operational controls, and human oversight provisions among its Annex A controls. The **EU AI Act** converts oversight from good practice into law. Article 4, in force since February 2025, requires providers and deployers to support the development of AI literacy among the staff who operate AI systems. Article 14 mandates effective human oversight of high-risk systems by competent natural persons, and Article 26(2) requires deployers to assign that oversight to persons with the necessary competence, training, and authority. Following the Digital Omnibus, these high-risk obligations apply from 2 December 2027 for standalone Annex III systems, and from 2 August 2028 for high-risk AI embedded in regulated products. Prudent programs are evidencing against them now, well ahead of the date.

Every one of these provisions terminates at a person. The framework can require oversight, but oversight is something a human does or fails to do under operational pressure. The organization that wrote a policy, ran a training, and logged the system has satisfied the easy two-thirds of the obligation. The remaining third, whether our people actually exercise the oversight the policy assumes, is the part that is currently undocumented, and it is the part an auditor will probe first.

Why Current Signals Do Not Qualify as Evidence

Organizations are not short of data about AI usage. They are short of data that an auditor will accept as evidence of human oversight. The distinction matters because the three artifacts most teams reach for each fail a basic evidentiary test: they do not observe the behavior the framework cares about.

Training completion records that a person was exposed to material. It says nothing about whether they can apply it. A professional can complete an AI-use course with a perfect quiz score and still accept a fabricated citation from a model the following week. Training completion answers “were they told?” when the framework asks “do they do it?”

Usage and activity logs record that AI was used and how often. Volume is not quality. A high-usage employee may be the one most likely to paste sensitive data into a prompt or to accept output uncritically. Activity dashboards show that work happened; they do not show that it was reliable. When a new system is introduced, execution is audited but behavior is not.

Self-assessment and attestation record what a person reports about their own practice. This is the weakest signal of the three, because confidence and competence diverge sharply in AI

collaboration. The population-level gap between how reliably people believe they verify AI output and how reliably they actually do is precisely where organizational risk concentrates. An attestation captures the belief, not the behavior, and the belief is systematically optimistic.

> *An auditor does not want to know what your people were told, how often they used the tool, or what they believe about themselves. The auditor wants to know what they do when the AI is wrong. Only observation answers that.*

What Behavioral Evidence Requires

If the qualifying signal is observed behavior, an evidence instrument for the People vector must meet four conditions. It must observe conduct rather than collect self-report. It must do so under realistic conditions, where the AI sometimes fails and the person must notice. It must produce a structured, comparable record that survives an audit. And it must generate that record without itself becoming a compliance liability: an instrument that observes professionals by hoarding their conversation content trades one governance problem for another.

PAICE is built to those four conditions. It is a **behavioral assessment** of human-AI collaboration: conversation is the medium through which behavior is observed, not the thing being scored. During a 15–25 minute session, PAICE injects errors, inconsistencies, and overconfident claims into AI responses (not trick questions, but replicas of the failures real AI systems produce) and observes whether the professional catches them. Catching an injected error is behavioral ground truth. Articulating a principle about verification is stated perception. Where the two conflict, the behavioral evidence dominates the score.

The assessment reports across five weighted dimensions, with Accountability (verification behavior, error detection, ownership of AI-informed decisions) carrying the highest weight precisely because it is the dimension most frameworks care about and the one most commonly underdeveloped.

Dimension	Weight	Behavioral focus relevant to governance
Accountability (A)	30%	Verification behavior: catching injected errors, checking outputs, owning AI-informed decisions. The core human-oversight signal.
Integrity (I)	25%	Recognizing AI limitations, maintaining boundaries, avoiding over-reliance on AI judgment.
Collaboration (C)	20%	Treating AI as a directed tool rather than an authority; effective iteration and context-setting.
Evolution (E)	15%	Adapting approach when AI behavior changes; learning within the session.

Dimension	Weight	Behavioral focus relevant to governance
Performance (P)	10%	Framing tasks clearly and driving toward outcomes.

Table 1: PAICE assessment dimensions. Accountability and Integrity, the two dimensions most directly tied to human-oversight obligations, together carry 55% of the composite weight.

Critically, PAICE meets the fourth condition by architecture. Conversation content is **not stored in production**: turn logging is disabled by code, not policy, and is auditable in the codebase. PII is redacted before any model sees a message. Participants in enterprise deployments are identified by opaque tokens (format XXXX-XXXX, roughly 1.1 trillion combinations) whose mapping to real employees lives in the organization's own systems, never in PAICE. Cohort analytics are returned only above a minimum of ten completed assessments. The instrument that generates your oversight evidence cannot itself become the next data-protection finding.

The Crosswalk: PAICE Output to Framework Requirements

The following table is the core of this paper. It maps, at a glance, the People-vector requirement in each major framework to the specific PAICE output that supplies evidence for it. It is designed to be lifted out of this document and placed directly into an audit file or a control-mapping spreadsheet. The detailed clause-by-clause mappings are in the Appendix. This crosswalk concerns AI-governance obligations; for the data-protection mapping (GDPR, CCPA, the NIST Privacy Framework) see the companion Verifiable Human-AI Collaboration paper noted in the Appendix.

Framework	People-vector requirement	PAICE evidence
NIST AI RMF 1.0	MEASURE: identify and apply metrics for the human-AI configuration and trustworthiness in context (MEASURE 2, 3, 4).	Dimensional behavioral scores per operator; aggregate cohort distribution as a recurring trustworthiness metric.
ISO/IEC 42001:2023	Competence of personnel (Cl. 7.2); human oversight and operational control (Annex A).	Objective competence evidence per role; periodic reassessment record for the AIMS audit trail.
EU AI Act Art. 14	Effective human oversight of high-risk AI systems by competent natural persons.	Behavioral demonstration that operators detect and act on AI error, scored and dated.
EU AI Act Art. 4	Sufficient AI literacy among staff who operate AI systems (in force Feb 2025).	Measured collaboration capability, not training attendance; evidence of literacy in practice.
EU AI Act Art. 26(2)	Deployer duty to assign oversight	Defensible record that assigned

Framework	People-vector requirement	PAICE evidence
	to persons with the necessary competence, training, and authority.	overseers meet a behavioral competence threshold.
Sector model-risk / professional-duty regimes	Effective challenge and competent supervision of model-assisted work (e.g. SR 11-7; professional competence rules).	Per-operator evidence of effective challenge: did the human override the model when it was wrong?

Table 2: PAICE-to-framework crosswalk (summary). Detailed clause-level mappings appear in the Appendix. The EU AI Act literacy duty (Art. 4) applies now; the high-risk human-oversight duties (Arts. 14, 26) apply from 2 December 2027 following the Digital Omnibus. PAICE supplies People-vector evidence; it does not certify conformity with any framework.

Placing PAICE Output Into an Audit File

Evidence is only useful if it can be presented. A PAICE deployment generates three artifacts a compliance officer can hand to an internal auditor, an external assessor, or a regulator, each addressing a different level of the governance question.

The first is the **individual assessment record**: a dated, dimensional score for a named role or opaque token, demonstrating that a specific overseer met a defined behavioral threshold before being assigned oversight duties. This is the artifact that answers Article 14 and Article 26 at the level of the individual in the loop.

The second is the **cohort readout**: an aggregate, privacy-preserving distribution across a team, function, or the whole organization, returned only above the ten-completion floor. This is the artifact that answers NIST AI RMF's MEASURE function and ISO/IEC 42001's expectation of a monitored, improving management system: it shows the trustworthiness of the human layer as a tracked metric over time, not a one-off snapshot.

The third is the **methodology and privacy record**: the documented scoring philosophy, the evidence hierarchy that places behavior above self-report, and the architectural privacy guarantees. This is what lets an assessor trust the first two artifacts: it establishes that the scores mean what they claim to mean and that generating them did not create a new data-protection exposure. Together the three move a compliance officer from "we have a policy on human oversight" to "we can show you the oversight happening, measured, and trending."

Honest Scope and Limitations

Overstating what an evidence instrument does is the fastest way to lose an auditor's confidence, so the boundaries of this claim should be explicit. PAICE supplies evidence for one vector of a governance program. It is not a substitute for the rest of it, and it is not a certification.

PAICE does not certify an organization against NIST AI RMF, ISO/IEC 42001, or the EU AI Act. Conformity with those frameworks is a whole-of-organization determination that includes system documentation, data governance, risk management process, and accredited audit where applicable. PAICE addresses the **human-oversight and competence dimension** specifically, and the mappings in this paper are PAICE's own analysis of where its output is relevant, not a representation by any standards body or regulator. Each organization's assessor makes the final call on what evidence satisfies a given control.

Two further limitations bear stating. First, a PAICE score reflects observed behavior during a bounded session; it is strong evidence of demonstrated capability, but like any assessment it is a sample, not a guarantee of conduct in every future interaction. Second, PAICE is calibrated against current-generation AI behavior, and its scoring baseline is maintained as models evolve; an organization relying on it for audit evidence should treat reassessment as periodic, in the same way competence evidence is refreshed elsewhere in a management system. These are the ordinary limits of behavioral measurement, and naming them is what makes the evidence credible rather than promotional.

Conclusion

The governance frameworks regulated industries now answer to have moved the evidence burden onto the behavior of the people operating AI. The system-level artifacts most programs already produce (documentation, logs, training records, attestations) do not reach that behavior, and auditors have begun to notice the gap. The People vector is the part of the obligation that is hardest to evidence and the part that is probed first.

PAICE was built to instrument exactly that vector: to observe how professionals behave when AI is wrong, to score it without retaining what was said or who said it, and to produce records that fit into an audit file. It does not certify conformity and does not replace the rest of a governance program. It supplies the one piece of evidence those programs are currently missing. For a compliance, risk, or security officer who has been asked to show that human oversight is real and not just policy, that is the artifact worth having.

About PAICE

PAICE.work PBC builds operational AI risk visibility through behavioral governance instrumentation. PAICE measures the human risk introduced into AI-enabled workflows (observed, not self-reported) and is the People-vector reference implementation within the AI Posture framework.

To map PAICE to your governance program, contact **info@paice.work** or visit paice.work.

Appendix: Clause-Level Framework Mappings

The summary crosswalk in Table 2 is expanded below, one framework per subsection. Each table lists the specific provision, what it asks for at the human layer, and the PAICE output that supplies evidence. These mappings are PAICE's analysis and are intended as a starting point for an organization's own control-mapping exercise, not as a conformity assertion.

A. NIST AI Risk Management Framework 1.0

NIST AI RMF is voluntary and outcome-oriented, organized into the Govern, Map, Measure, and Manage functions. PAICE output is most directly relevant to the Measure function, where the human-AI configuration is named as a subject of measurement.

Function / subcategory	Human-layer expectation	PAICE evidence
GOVERN: roles & competence	Accountability structures ensure personnel are trained and equipped for their AI-related duties.	Behavioral baseline per role; reassessment cadence as a governance control.
MEASURE 2: trustworthiness	Apply metrics for system trustworthiness in the deployment context, including human factors.	Cohort score distribution as a recurring human-reliability metric.
MEASURE 3: tracking	Mechanisms to track identified risks over time.	Period-over-period cohort trend; movement after interventions.
MEASURE 4: feedback	Feedback on efficacy of measurement, including from operators in context.	Per-operator catch/miss patterns surfaced as actionable signal.
MANAGE 1: prioritization	Risks prioritized and acted upon based on assessment.	Identifies where the human layer is weakest, informing targeted action.

Table A1: NIST AI RMF mapping. PAICE primarily evidences the MEASURE function for the human-AI configuration.

B. ISO/IEC 42001:2023

ISO/IEC 42001 is the certifiable AI management system standard. Its main clauses establish the management system; Annex A provides reference controls. PAICE evidence supports the competence and human-oversight elements.

Clause / control	Human-layer expectation	PAICE evidence
Cl. 7.2: Competence	Determine and ensure competence of persons affecting AI performance; retain evidence.	Objective, dated competence record per role; retainable evidence for the AIMS.
Cl. 9.1: Monitoring	Monitor and measure the AI management system's performance.	Recurring cohort measurement of the human layer over time.
Cl. 10: Improvement	Continually improve the suitability and effectiveness of the AIMS.	Trend data showing the effect of training or process changes on behavior.
Annex A: human oversight	Controls for human oversight of AI systems and responsible use.	Demonstrated oversight capability of assigned personnel.

Table A2: ISO/IEC 42001 mapping. PAICE evidences Clause 7.2 competence and the monitoring / improvement loop for the human layer.

C. EU AI Act

The EU AI Act is binding law (in force since August 2024). The provisions below concern the human operators of AI systems rather than the systems themselves, and PAICE output speaks directly to them. The Article 4 literacy duty has applied since February 2025 and, following the Digital Omnibus, obliges actors to support the development of staff AI literacy. The high-risk human-oversight duties in Articles 14 and 26 apply from 2 December 2027 for standalone Annex III systems, and from 2 August 2028 for high-risk AI embedded in regulated products. Whether a given system is high-risk, and which obligations attach, is a legal determination for the organization.

Article	Human-layer obligation	PAICE evidence
Art. 4: AI literacy	Support the development of AI literacy among staff operating AI systems (in force since Feb 2025, post-Omnibus wording).	Measured collaboration capability in practice, beyond training attendance.
Art. 14(4)(d): override	Overseer of a high-risk system must be able to decide not to use, disregard, override, or reverse its output.	Direct behavioral evidence: did the operator override the model when it was wrong?
Art. 26(2): deployer duties	Assign oversight to natural persons with the necessary competence, training, and authority.	Defensible record that the assigned person met a behavioral competence

Article	Human-layer obligation	PAICE evidence
		threshold.

Table A3: EU AI Act mapping. PAICE evidences the operator-competence and human-oversight obligations; it does not address provider-side system obligations.

D. Sector regimes (illustrative)

Beyond the cross-sector frameworks, regulated industries carry domain-specific duties that increasingly extend to AI-assisted work. The examples below are illustrative of how PAICE evidence maps; the controlling text in each case is the applicable regulation or professional rule.

Domain	Representative duty	PAICE evidence
Financial services	Model risk management and effective challenge of model-assisted decisions (e.g. SR 11-7 principles).	Per-operator evidence of effective challenge: overriding the model when it errs.
Legal practice	Duty of technological competence and supervision of work product.	Behavioral evidence that the practitioner verifies AI output before relying on it.
Healthcare	Clinician responsibility for decisions informed by AI-enabled tools.	Demonstrated verification behavior and appropriate non-reliance on AI judgment.

Table A4: Illustrative sector mappings. PAICE supplies operator-level behavioral evidence; the controlling obligation is set by each sector's regulator or professional body.

E. Related PAICE portfolio papers

PAICE's portfolio papers divide along three axes: privacy/verifiability (is the assessment data handled lawfully?), measurement science (are the scores defensible?), and AI governance (can we evidence human oversight?). A compliance officer typically routes the privacy paper to the DPO or privacy counsel, the measurement paper to internal audit or the model-risk function, and this paper to the AI-governance owner. The two newest portfolio briefs (Aggregated Intelligence and One Number You Can Defend, June 2026) sit above all three, framing why an independent People-vector measure matters at the leadership level. The table below shows what each carries and how current its regulatory content is. Canonical index: paice.foundation/papers/.

Paper	Focus	Regulatory mapping it carries
Aggregated Intelligence (exec brief, June 2026)	The case for measuring human-machine collaboration independently while the	Not a regulatory mapping; establishes why an independent measure of the people-AI layer

Paper	Focus	Regulatory mapping it carries
	bottleneck is still visible. Investor and board framing.	is load-bearing.
One Number You Can Defend (exec brief, June 2026)	Details the AI Posture instrument: one governance score, bounded by the weakest vector. For GRC leaders. Companion to Aggregated Intelligence.	Orthogonal to specific frameworks: a composite governance score derived from PAICE and sibling vector measures.
Verifiable Human-AI Collaboration (NEARCON 2026, Feb 2026)	Privacy-preserving assessment with cryptographic integrity: behavioral observation routed through TEE-protected inference; on-chain score attestation.	GDPR, CCPA/CPRA, NIST Privacy Framework, plus TEE / on-chain attestation controls. Predates recent EU AI Act developments covered here.
Closing the Collaboration Gap (ISPI 2026, March 2026)	Maps PAICE measurement to established human-performance-technology frameworks (Gilbert, Mager, Rummler-Brache, Thalheimer, Phillips, Brinkerhoff).	Not a regulatory mapping; establishes the behavioral-skill foundation for why the scores are defensible evidence.
PAICE Framework v4 (whitepaper, Nov 2025)	Rationale, framework, and architecture behind PAICE as the evidence layer for AI governance; measures observable behavior rather than self-report.	Background and architecture; not a clause-level regulatory mapping.
The People-Vector Evidence Layer (this paper, June 2026)	Behavioral output as evidence for the human-oversight vector of AI governance.	NIST AI RMF 1.0, ISO/IEC 42001:2023, EU AI Act (Arts. 4, 14, 26). Current authority for the AI-governance mapping.

Table A5: PAICE portfolio papers (source: paice.foundation/papers/papers.json). The privacy/verifiability, measurement-science, and AI-governance papers map PAICE to three different reviewer audiences; where they overlap on currency, this paper carries the more recent regulatory position.

AUTHOR VALIDATION CHECKLIST: REMOVE BEFORE PUBLICATION

This page is an author tool. It is not part of the whitepaper and must be deleted before the document is distributed. Verify every item below against current source documents before publishing.

Framework claims, verified 2026-06-06 (every-ai-law data/instruments/, last_verified 2026-06-02; Perplexity sonar-pro cross-check)

[x] NIST AI RMF 1.0 functions correct (Govern, Map, Measure, Manage; every-ai-law data/instruments/nist-ai-rmf.md, last_verified 2026-03-25). MEASURE 2/3/4 and MANAGE 1 confirmed substantively. GOVERN pin de-numbered to function level (GOVERN 3 is workforce-diversity, not roles/competence; corrected to avoid wrong pinpoint).

[x] ISO/IEC 42001:2023 clauses confirmed (every-ai-law data/instruments/iso-42001.md, last_verified 2026-03-26): 7.2 Competence (retain documented evidence), 9.1 Monitoring/measurement/analysis/evaluation, 10 Improvement; Annex A responsible-use / human-oversight controls. Publication 2023-12-18.

[x] EDITORIAL ASSUMPTION (author direction): the Digital Omnibus is treated as valid and in force throughout this paper. Stated as settled: Art. 4 softened from ensure to support (in force since 2025-02-02); standalone Annex III high-risk duties (Arts. 14, 26) apply 2 December 2027; embedded high-risk 2 August 2028; Annex III scope (employment) not narrowed. Provenance: every-ai-law data/instruments/eu-ai-act.md amendment entry (last_verified 2026-06-02). Art. 26(2) / Art. 14(4)(d) pinpoints per Reg. 2024/1689.

[x] SR 11-7 / OCC 2011-12: 'effective challenge' confirmed as defined term; framed as illustrative.

[x] No claim that PAICE certifies conformity with any framework anywhere in the body.

[x] Companion-paper titles (Table A5) reconciled against paice.foundation/papers/papers.json (last_updated 2026-06-06): Aggregated Intelligence (June 2026), One Number You Can Defend (June 2026), Verifiable Human-AI Collaboration (NEARCON 2026), Closing the Collaboration Gap (ISPI 2026), PAICE Framework v4 (Nov 2025). Replaces prior duplicate 'Privacy, Security and Compliance' row (same paper as Verifiable Human-AI Collaboration) and prior misnamed 'Measuring Human-AI Performance' row (actual title: Closing the Collaboration Gap).

[x] Omnibus treated as adopted for purposes of this paper; no OJ-publication caveat carried in body prose.

PAICE product facts to verify (re-check against source docs)

[ ] Five dimensions and weights correct: A 30 / I 25 / C 20 / E 15 / P 10 (source: PRD.md).

[ ] 0–1000 scale and five tiers accurate (source: SCORING_PHILOSOPHY_AND_DESIGN_INTENT.md).

[ ] should_log_turns still False in production; no conversation content stored (source: PRODUCTION_PRIVACY_POLICY.md).

[ ] PII redaction before LLM still active (source: PII_SERVICE.md).

[ ] Opaque token format XXXX-XXXX, ~1.1 trillion combinations, 32-char alphabet (source: PILOT_PROGRAM_PRIVACY_ARCHITECTURE.md).

[ ] Cohort aggregation minimum still 10 completions.

[ ] Session length 15–25 minutes still accurate.

[ ] Brand rule: P = People + AI Collaboration Effectiveness (never “Professional”).

Positioning / language

[ ] Category language matches POSITIONING_CANON.md (operational AI risk visibility through behavioral governance instrumentation).

[ ] No banned phrasing (“master,” “personalized insights,” “tips and best practices,” “AI literacy” used only in contrast).

[ ] AI Posture / People-vector framing consistent with INTENT.md and portfolio scope.

[ ] No em dashes in body prose; URLs bare https; copyright PAICE.work PBC throughout.

Document integrity

[ ] PDF page count reasonable; no blank pages; tables render with borders and padding.

[ ] Headers/footers on body pages only, not title page; page numbers at correct size.

[ ] This checklist section removed before publication.