Governance Without Surveillance

Why Privacy by Architecture Is What Makes Behavioral AI Measurement Adoptable

Sam Rogers, Founder of PAICE.work PBC

June 2026 • Version 1.0

For works councils, employee representatives, labor counsel, and data protection officers

Executive Summary

Any tool that measures how employees work with AI invites one objection before any other: is this surveillance? In a European context, with a works council across the table, that objection is not a public-relations concern. It is a gate. A technical system that can monitor employee behavior triggers mandatory co-determination, employee-information duties, and data-protection scrutiny, and a measurement programme that cannot answer the surveillance question dies in consultation regardless of how useful its output would have been.

This paper makes the case that the objection is answerable, but only by architecture. Policy promises ("we won't misuse the data") do not survive a works council, a data protection officer, or a change of corporate ownership, because they protect data that still exists and can still be re-pointed at an individual. **PAICE (People + AI Collaboration Effectiveness)** is built so the data that would make it surveillance is never created: no user accounts, identity reduced to an irreversible hash, conversation content never written to the production store, and enterprise participants represented by opaque tokens whose mapping to real employees lives only in the organization's own systems.

The consequence reframes privacy entirely. In most assessment tools privacy is a defensive feature, something added to reduce breach risk. In PAICE it is the enabler: the architecture is the reason a works council can approve the deployment, the reason a DPO can sign the data-protection assessment, and the reason employees can be told the truth about what is and is not retained. Governance of AI-assisted work becomes possible precisely because the measurement is built to not be surveillance. This paper defines surveillance structurally, shows where PAICE fails that definition by design, and maps the architecture to the European legal touchpoints (works-council co-determination, the EU AI Act workplace-information duty, and GDPR data minimisation) that decide whether a programme is adoptable.

This is one paper in PAICE's whitepaper series. Where the companion **Privacy, Security and Compliance** paper documents the data-protection controls and **The People-Vector Evidence Layer** maps PAICE to AI-governance frameworks, this paper addresses the adoption question that precedes both: whether employees and their representatives will accept the measurement at all. The Appendix lists the companion papers.

The Objection That Stops Deployment

A programme to measure AI-collaboration behavior usually fails before it is evaluated on its merits. The failure happens in the room where it must be approved: with the works council, the employee representative body, the data protection officer, or labor counsel. Their first question is not "is the measurement accurate?" It is "what does this do to the people being measured?" If

the honest answer is “it records what each named employee does and keeps it,” the conversation is effectively over in any jurisdiction with strong worker protections.

This is not caution for its own sake. In Germany, for example, the works council holds a co-determination right under **Section 87(1) no. 6 of the Works Constitution Act (BetrVG)** over the introduction and use of technical devices designed to monitor employee behavior or performance. Settled Federal Labour Court case law has read that right broadly: the trigger is the device’s **objective capability** to capture behavior- or performance-related data, not the employer’s intention to monitor. Practically any IT system that can record what an identifiable employee did can fall within it. A tool that scores individual behavior and retains it is squarely inside that test.

So the adoption barrier is not really about measurement. It is about the by-products of measurement: identity, retention, and the latent ability to use the record against the person later. Most tools accept those by-products as the price of insight and then try to manage them with policy. That is the move that fails. The rest of this paper argues for a different one: design the by-products out, so the measurement can proceed.

What Surveillance Actually Is

To answer the objection precisely, define the thing precisely. Observation is not surveillance. A coach watching a trainee, an examiner scoring a test, an aggregate productivity metric: none of these is inherently surveillance. Surveillance is observation that carries three additional properties, and it is the combination, not the observation, that worker-protection law and employee intuition both react to.

First, **identification**: the observation is tied to a known individual. Second, **retention**: a durable record of that individual’s conduct is kept. Third, **repurposability**: the retained, identified record can later be turned to a use the person did not agree to and cannot see, from performance management to discipline to litigation. A system with all three is surveillance whatever its stated purpose. Remove any one and the character changes; remove all three and what remains is measurement.

> *Surveillance is not observation. It is observation tied to an identity, retained as a durable record, and reusable against the person later. Strip those three properties and measurement is what is left.*

This definition is useful because it is testable. Instead of arguing about intentions, a works council or DPO can ask three architectural questions: Can the system attribute this record to a named person? Is the record kept? Could it be repurposed against them? A tool that must answer yes to all three cannot escape the surveillance objection through assurances. A tool that answers no by construction has a different conversation available to it.

How PAICE Fails the Surveillance Test by Design

PAICE is engineered to answer no to all three questions, and to do so structurally rather than by policy. Each property of surveillance is removed at the level of what data the system creates and keeps, which is why the answers hold even against a hostile party or a future owner.

No identification

PAICE requires no user accounts: no registration, username, password, or persistent profile. A user is reduced to a SHA-256 hash that cannot be reversed. In enterprise deployments, participants are represented by opaque tokens of the form XXXX-XXXX, randomly drawn from a 32-character alphabet (roughly 1.1 trillion combinations) and carrying no encoded identity. The mapping from token to employee is held by the organization in its own systems and is never given to PAICE. PAICE cannot attribute a score to a named person, even from its own database, because it never holds the link.

No retention of the observed material

The conversation through which behavior is observed is not stored in the production environment. Turn-by-turn logging is disabled by code, not by configuration or promise, and is auditable in the codebase. During a session the transcript exists in memory to drive scoring; once scores are computed, the raw material is not written to any persistent store. Before any text reaches a model, a redaction step removes emails, phone numbers, government identifiers, and similar values, so even the transient processing does not handle the sensitive originals. What persists is the score, not the conduct that produced it.

No repurposability

Because there is no identified, retained record of individual conduct, there is nothing to repurpose. Organizational analytics are returned only as aggregates above an enforced floor of ten completed assessments, which prevents a small-group readout from being decomposed back to an individual. A score detached from identity and from the underlying conversation cannot be turned into a disciplinary exhibit, because it neither names the person nor preserves what they did. The latent threat that makes employees distrust monitoring tools is removed at the source.

Property of surveillance	The question a works council asks	PAICE answer, by architecture
Identification	Can a record be tied to a named employee?	No. Irreversible hash; opaque tokens; identity map held only by the employer, never by PAICE.
Retention	Is the conduct kept?	No. Conversation content not stored in production (disabled in code); PII redacted before processing. Only scores persist.
Repurposability	Could the record be used against the person later?	No. Nothing identified or retained to repurpose; analytics only above a 10-completion aggregation floor.

Table 1: The surveillance test applied to PAICE. The architecture answers no to each property, which is what distinguishes measurement from monitoring.

Why the Architecture Pre-Answers the Law

Removing the three properties is not only ethically cleaner; it is what makes a deployment approvable under the specific legal instruments that govern measurement of employees in Europe. The architecture does the work that policy language is asked, and usually fails, to do.

Works-council co-determination

The German co-determination test under **BetrVG § 87(1) no. 6** turns on whether a technical device is objectively capable of monitoring an identifiable employee's behavior or performance. PAICE does not eliminate the works council's role, and should not be presented as doing so, but it changes what the council is asked to approve. Instead of negotiating safeguards around a system that holds identified behavioral records, the council evaluates a system that holds none. The consultation starts from "there is no individual record to misuse" rather than "trust the retention policy," which is a categorically easier approval to grant.

EU AI Act workplace information duty

Where an AI system used in employment is high-risk, the EU AI Act imposes a direct duty on the employer. **Article 26(7)** requires that, before putting such a system into use at the workplace, deployers who are employers inform workers' representatives and the affected workers that they will be subject to it. PAICE's architecture makes that disclosure straightforward and credible: the employer can tell workers exactly what is and is not retained, and the statement survives scrutiny because it is enforced in code. A transparency duty is far easier to discharge honestly when the honest description is reassuring. (Article 26(7)'s substance is unchanged by the Digital Omnibus, but it attaches with the high-risk regime, which

applies to standalone Annex III systems from 2 December 2027. Transparency with workers is sound practice to adopt well ahead of that date.)

GDPR data minimisation

GDPR's data-minimisation principle, **Article 5(1)(c)**, requires that personal data be adequate, relevant, and limited to what is necessary for the purpose. A measurement tool that retains identified conversation transcripts has a minimisation problem to argue away. PAICE's design is minimisation taken to its conclusion: the necessary output is a score, so the conversation and the identity are not retained at all. **Article 88** further lets member states and collective agreements set stricter employment-context rules, which is exactly the terrain a works council operates on; an architecture that already minimises gives the DPO and the council less to litigate and more to approve.

Privacy as Adoption Enabler, Not Security Feature

Most vendors present privacy as risk reduction: encryption, access controls, retention limits, all aimed at making a breach less likely or less damaging. Those controls matter and PAICE implements them, but they are defensive, and they operate on data the system has chosen to hold. Framed that way, privacy is a cost center that competes with the product's usefulness.

The argument of this paper is that for measurement of people, privacy is the opposite of a cost: it is the precondition for the product existing in the workplace at all. An organization cannot measure AI-collaboration behavior across its workforce if the measurement is surveillance, because the workforce's representatives will not permit it. By removing identity, retention, and repurposability at the architectural level, PAICE converts an un-approvable programme into an approvable one. The privacy architecture is not protecting the data; it is unlocking the use case.

> *In a monitoring tool, privacy reduces the harm of the data you keep. In PAICE, privacy is why there is no such data, and why the measurement is allowed to happen. The architecture is the adoption strategy.*

This is also why "governance without surveillance" is not a slogan but a structural claim. Organizations genuinely need visibility into how their people operate under AI-assisted conditions; that need is real and growing. The false assumption is that obtaining it requires watching individuals. PAICE demonstrates the assumption is false: the visibility an organization needs is behavioral and aggregate, and it can be produced without the identified, retained record that turns visibility into surveillance.

Honest Scope and What the Employer Still Must Do

The architecture makes a deployment adoptable; it does not make the employer's obligations disappear, and presenting it that way would rightly cost credibility with the very audiences this paper addresses. Three boundaries should be explicit.

First, PAICE does retain some data, and it should be described honestly. It stores assessment scores and metadata (timestamps, turn counts), an irreversible user hash, and, where a participant chooses to provide an email, that address in encrypted form. None of this is identified behavioral content, but it is not **nothing**, and a works council or DPO is entitled to the full list rather than a reassuring summary.

Second, the architecture supports the legal process; it does not replace it. A works-council consultation, an employee-information notice under EU AI Act Article 26(7) where applicable, and a data-protection assessment may all still be required. PAICE's contribution is that each of these becomes easier to complete and likelier to succeed, because the honest description of the system is favorable. The employer must still do them.

Third, this paper is not legal advice, and the mappings are PAICE's own analysis of where its architecture is relevant, not a determination by any regulator, court, or works council. Whether a particular deployment is high-risk under the EU AI Act, or triggers co-determination in a given jurisdiction, is a legal question for the organization and its advisors. What the architecture guarantees is narrower and more durable than a legal conclusion: there is no identified, retained record of individual conduct for anyone to misuse.

Conclusion

The objection that stops behavioral AI measurement in regulated and worker-protected environments is the surveillance objection, and it is usually correct, because most measurement tools are built in a way that cannot answer it. The answer is not a better privacy policy. It is an architecture in which the data that would make the tool surveillance is never created: no identities, scores only, no conversation logging.

That design choice is what makes the difference between a programme that dies in consultation and one a works council can approve. It is why privacy, for a tool that measures people, is not a defensive feature bolted on at the end but the thing that makes the use case possible in the first place. Governance without surveillance is not a softening of the goal; it is the only version of the goal that is adoptable. PAICE is built to be that version.

About PAICE

PAICE.work PBC builds operational AI risk visibility through behavioral governance instrumentation. It measures the human risk introduced into AI-enabled workflows without identities, without conversation logging, and without an individual record that could be repurposed against the person measured.

For a works-council or DPO review pack, contact **info@paice.work** or visit paice.work.

This document is provided for informational purposes only and does not constitute legal, compliance, or labor-relations advice. References to the German Works Constitution Act (BetrVG § 87(1) no. 6), the EU AI Act (Regulation (EU) 2024/1689, Article 26(7)), and the GDPR (Articles 5(1)(c) and 88) describe those instruments as published and reflect PAICE.work PBC's own analysis of where its architecture is relevant; they are not endorsed by any regulator, court, or worker-representative body. Whether a deployment triggers co-determination, is high-risk under the EU AI Act, or requires a data-protection assessment is a determination for the organization and its advisors.

Appendix A: European Legal Touchpoints

The provisions below are the instruments most often raised when measurement of employees is proposed in Europe. Each is mapped to the PAICE architectural feature relevant to it. Citations were verified against primary sources; pinpoints and the EU AI Act timing reflect the position as of May 2026 and are subject to revision.

Provision	What it requires (human / employee layer)	Relevant PAICE feature
BetrVG § 87(1) no. 6 (Germany)	Works-council co-determination over technical devices objectively capable of monitoring employee behavior or performance (capability test per BAG case law, not intent).	No identified behavioral record exists to monitor; consultation starts from architecture, not policy.
EU AI Act Art. 26(7)	Employer-deployers must inform workers' representatives and affected workers before workplace use of a high-risk AI system. (Substance unchanged by the Omnibus; high-risk regime applies from 2 December 2027.)	Honest, code-enforced description of what is and is not retained makes the disclosure credible.
GDPR Art. 5(1)(c)	Data minimisation: personal data adequate, relevant, and limited to what is necessary for the purpose.	Only scores persist; conversation and identity are not retained. Minimisation by construction.
GDPR Art. 88	Member states / collective agreements may set stricter employment-context processing rules.	An already-minimised architecture leaves less for stricter local rules to constrain.

Table A1: European legal touchpoints. PAICE supplies architecture that eases each obligation; the controlling determination remains with the organization and its advisors.

Appendix B: Related PAICE whitepapers

PAICE's whitepapers address adjacent questions for adjacent readers. This paper answers "will employees and their representatives accept the measurement?" Route it to works councils, employee representatives, labor counsel, and DPOs. The companions answer the data-protection, AI-governance, and measurement-science questions.

Paper	Question it answers / focus	Primary reader
Governance Without Surveillance (this paper)	Will the workforce accept the measurement? Privacy architecture as adoption enabler.	Works councils, employee reps, labor counsel, DPOs.
Privacy, Security and Compliance	Is the data handled lawfully and securely? GDPR, CCPA, NIST Privacy Framework.	DPO, privacy counsel, security review.
The People-Vector Evidence Layer	Can we evidence human oversight? NIST AI RMF, ISO/IEC 42001, EU AI Act.	Risk / compliance / AI-governance owner.
Measuring Human-AI Performance (ISPI 2026)	Why are the scores defensible? Measurement science and validity.	Research, assessment, methodology reviewers.
Verifiable Human-AI Collaboration (NEARCON 2026)	Can privacy and integrity be proven, not just promised? TEE inference, on-chain attestation.	Technical security, procurement.
AI Posture (umbrella framework, aiposture.org)	What is our overall AI readiness? AIP across People, Infrastructure, and Regulation; this paper is the People-vector adoption case.	Board / GRC leadership (portfolio level).

Table B1: PAICE whitepaper series. This paper is the adoption / employee-trust entry; the surveillance argument here complements the data-protection detail in the Privacy paper.