Evidence Projection: How One Authorization Boundary Satisfies Many Jurisdictional Regimes

One enforcement reality. Many evidentiary presentations.

Edward Meyman
FERZ, Inc., McLean, Virginia

May 2026 · Version 1.0.0 · CC BY 4.0 · doi.org/10.5281/zenodo.20277842

ABSTRACT

AI compliance is fragmenting along jurisdictional lines. AI execution is not. Enterprises operating across the European Union, multiple US states, sectoral regulators, and Chinese algorithm-registration and AI governance regimes face partially overlapping, partially conflicting evidentiary obligations. The conventional response, parallel compliance architectures with separate governance stacks per jurisdiction, fails not only economically but structurally: parallel stacks govern the same underlying action separately and produce documentation that diverges from operative enforcement over time. This article argues that the regulatory convergence underneath fragmentation moves in one direction. Across regimes, regulators are increasingly orienting their evidentiary expectations toward execution-time enforcement rather than after-the-fact documentation. Once an organization can produce a reconstructable governance verdict bound to each execution event, jurisdictional fragmentation reduces from a governance-architecture problem to an evidence-format problem. The article develops the projection model, in which a single execution-time authorization boundary produces proof-carrying decision artifacts that project into jurisdiction-specific evidence formats. The result is one enforcement reality and many evidentiary presentations.

KEYWORDS

AI governance, AI compliance, execution-time authorization, runtime authorization, evidence projection, proof-carrying decision artifacts, jurisdictional AI compliance, Four Tests Standard, fail-closed AI governance, AI authorization boundary.

AI compliance is fragmenting along jurisdictional lines. AI execution is not.

The next regulatory cycle will require enterprises operating across the European Union, multiple US states, sectoral regulators, and Chinese algorithm-registration and AI governance regimes to satisfy simultaneous, partially overlapping, partially conflicting evidentiary obligations from each. The conventional response is to build parallel compliance architectures: separate governance stacks, separate regional workflows, separate documentation pipelines, separate oversight systems.

That response does not scale economically or operationally. It also fails on a deeper structural axis.

The convergence happening underneath jurisdictional fragmentation moves in one direction. Across regimes, regulators are increasingly orienting their evidentiary expectations toward execution-time enforcement rather than after-the-fact documentation. Once an organization can produce a reconstructable governance verdict bound to each execution event, jurisdictional fragmentation reduces from a

governance-architecture problem to an evidence-format problem. The same authorization artifact can be projected into the evidentiary surface each regime requires.

Why parallel compliance stacks fail structurally

The economic argument against parallel stacks is the surface argument. Costs scale with the number of regimes. Engineering, legal, and operations duplicate work across architectures that do substantially similar things in incompatible ways. That argument is real but it is not the deepest one.

The structural failure is sharper. Each parallel stack governs the same underlying AI action separately. A single execution event in production cannot be governed by multiple authorization boundaries operating in parallel without one boundary dominating in practice, or the boundaries fighting and one of them being overridden. Whatever actually authorizes the action becomes the operative governance event. Everything else is documentation of that event reformatted per jurisdiction.

Parallel governance stacks therefore always degenerate into a primary enforcement layer plus N parallel documentation layers. The documentation layers diverge from the enforcement layer in ways that compound over time: control descriptions drift, control coverage diverges, oversight assumptions diverge. When a regulator in any single jurisdiction asks what actually happened at execution time, the documentation in their format may not reconstruct it. Discrepancies between what each parallel stack documents and what actually executed create evidentiary risk in every regime simultaneously.

Parallel stacks treat governance as a documentation problem. That treatment works only when the evidentiary unit is the document. Once the evidentiary unit becomes the execution event itself, parallel documentation cannot substitute for a unified authorization decision. The architecture inverts.

The convergence vector

Regulatory regimes that differ substantially on substantive scope are converging on a shared structural question: not what your policy says, but what was authorized at execution time, and on what basis.

This direction is visible across several regimes. Conformity expectations under European AI legislation push toward demonstrating that controls applied to actual operation, not merely that controls existed on paper. US sectoral and state-level approaches in higher-risk contexts increasingly request execution-time evidence rather than policy attestation. Inspection regimes in other jurisdictions are oriented toward reconstructable decisions and reviewable bases for AI-driven actions.

The claim is not that these regimes use identical regulatory language. They do not. The claim is that their evidentiary direction converges on the same structural requirement: a reconstructable governance verdict bound to a specific execution event. The artifact must carry sufficient internal structure to show which policy boundary applied, which controls were active, whether oversight conditions were satisfied, what risk thresholds were evaluated, which verdict (ALLOW, DENY, or ABSTAIN) the authorization boundary issued and on what basis, and whether the decision can be independently reconstructed.

That artifact is qualitatively different from a dashboard, an approval queue, a guardrail intervention, or a post-hoc audit trail.[5] It is bound to the execution event. It carries its own proof. It is replayable. Prior work establishes the architectural necessity of this artifact class and the analytical reasons monitoring and

documentation cannot substitute for it.[1] [2] [3]

The Evidence Projection Model

Define an evidentiary projection as a transformation from a single authorization decision artifact into a regime-specific evidence format. The underlying authorization artifact is the source. The jurisdictional evidence package is the projection.

Projection does not require the authorization artifact to reconstruct every internal model state, latent inference, or planning step. The governed unit is the point at which an AI-mediated process attempts to bind to external action: a tool call, transaction, disclosure, workflow transition, approval, deployment, or other operational consequence. Long-running and agentic systems are therefore governed as ordered sequences of authorization events, each capable of producing its own artifact or sub-artifact, rather than as one opaque cognitive episode. Each authorization event resolves to one of three verdicts, ALLOW, DENY, or ABSTAIN, with fail-closed semantics on ABSTAIN.

For projection to work, the underlying authorization artifact must contain enough information to support every required projection. This places specific structural requirements on the artifact.

It must be **reproducible**. Given the inputs and the policy version in force at decision time, anyone with the artifact can reproduce the verdict. The decision is not an opaque output of an opaque process.

It must be **verifiable**. The artifact binds the decision to the specific policy version applied, the governed state at decision time, and the inputs evaluated. Provenance is recoverable from the artifact itself, not reconstructed by interpretation after the fact.

It must be **complete** relative to anticipated projections. The artifact contains the information each jurisdictional projection requires. Anything any projection might need is present in the source, so no projection requires fabrication or reinterpretation.

It must be **bounded**. The artifact does not encompass more than the authorization decision actually established. Projections cannot extract claims the underlying decision did not support, and the boundary between what was authorized and what was not remains intact in every derived projection.

These properties are not informal aspirations. They are testable, and the open vendor-neutral Four Tests Standard (STOP, OWNERSHIP, REPLAY, ESCALATION) specifies testable criteria for authorization artifacts of this class.[4]

The four tests are not arbitrary. They verify integrity properties that the three-verdict space requires. STOP tests whether DENY and ABSTAIN verdicts actually halt execution. OWNERSHIP tests whether any verdict can be traced to its policy version and governing entity. REPLAY tests whether any verdict can be deterministically reproduced from its recorded inputs and policy version. ESCALATION tests the human-override pathway that ABSTAIN requires.

An authorization boundary that produces artifacts satisfying these properties is projectable. From a single decision artifact, separate projection mappings can produce EU-format evidence, US sectoral evidence, state-law evidence, and inspection-regime evidence, each adapted to the format the regime requires, each

derived from the same source decision. The projections may differ in vocabulary, emphasis, and packaging. They cannot differ on what was authorized, because they all derive from the same authorization event.

Projection also does not mean every jurisdiction asks the same legal question or accepts the same filing format. It means the same execution-time enforcement record supplies the evidentiary substrate from which jurisdiction-specific showings can be generated without inventing parallel enforcement realities.

This is the architectural inversion. Parallel governance stacks attempt to operate N enforcement systems in parallel and produce N documentation streams from N parallel realities. Projection replaces parallel enforcement realities with one shared enforcement reality and many evidentiary presentations.

One enforcement reality. Many evidentiary presentations.

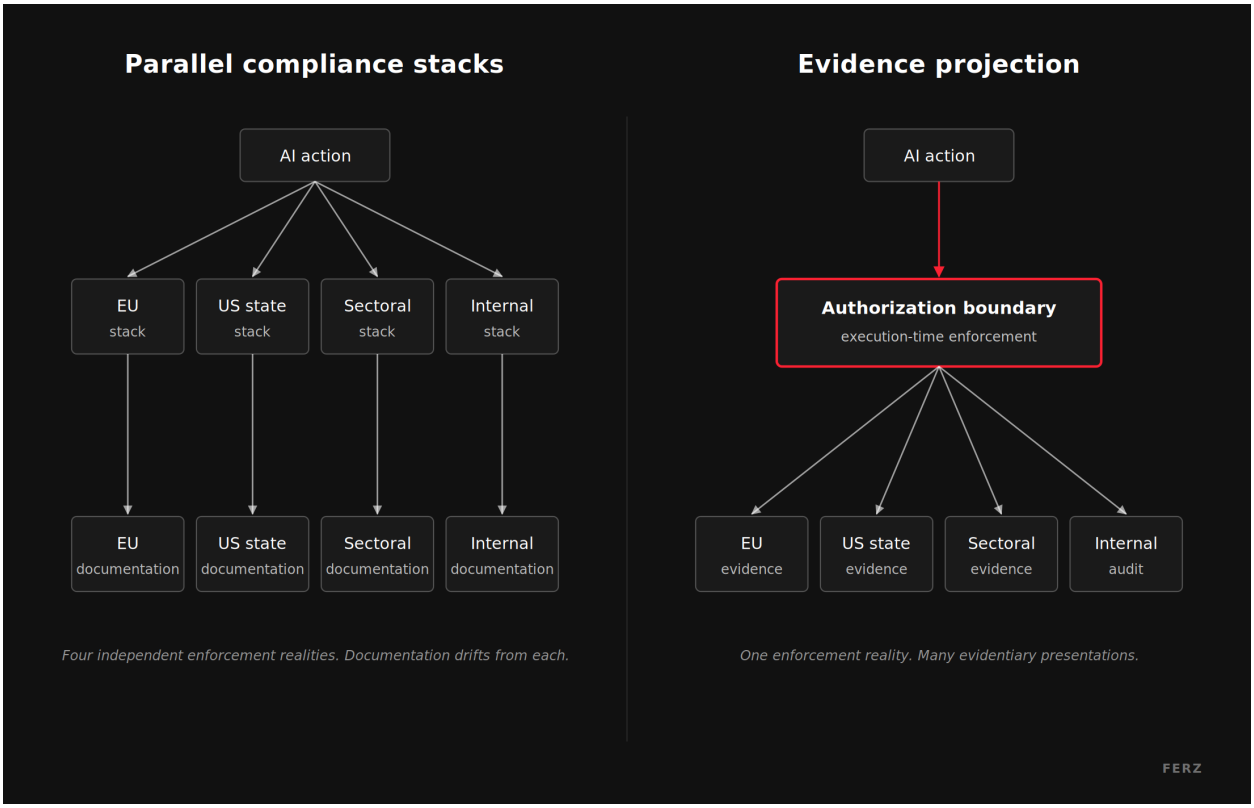

Figure 1. *Parallel compliance stacks contrasted with evidence projection. Left panel: under a parallel-stack architecture, a single AI action is governed independently by N jurisdictional stacks (EU, US state, sectoral, internal in this illustration), each producing its own documentation. Because each stack operates as a separate authority over the same underlying action, documentation diverges from operative enforcement over time and across regimes. Right panel: under evidence projection, the same AI action passes through one execution-time authorization boundary (red), which records a single decision artifact. Jurisdiction-specific evidence formats are derived from that artifact as projections. The number of regimes does not change the number of enforcement events, only the number of evidentiary presentations derived from one shared enforcement reality.*

An illustration

Consider an AI-assisted denial or approval of a medical coverage request in a US health insurance context. At execution time, the authorization boundary evaluates the request against the active policy version, the governed state of the applicable clinical and coverage rules, and the inputs presented. The verdict, ALLOW, DENY, or ABSTAIN, is recorded as a single authorization artifact. The artifact carries the policy version reference, the inputs evaluated, the clinical and coverage rules cited, the risk score and applicable threshold at evaluation time, the basis for the verdict, and cryptographic provenance information sufficient for independent verification of the decision. An ABSTAIN verdict halts execution and routes the decision to authorized human override rather than auto-approving or auto-denying under ambiguous conditions.

That single artifact supports multiple projections from one source:

- An EU AI Act conformity projection demonstrating that controls applicable to a high-risk system operated as defined at the moment of the decision.
- A US state-law projection demonstrating compliance with applicable AI-decisioning and utilization-review requirements.
- A HIPAA or sectoral health-regulation projection demonstrating the role of protected information in the decision and the scope of its use.
- An internal audit projection supporting risk-committee review with consistent provenance.

Each projection presents the same authorization event in the vocabulary, scope, and emphasis its recipient requires. None adds facts the underlying decision did not establish. None disagrees with the others about the verdict or its basis.

Implications for enterprise governance architecture

The projection model changes the strategic question facing enterprises subject to multi-jurisdictional AI governance.

The question is no longer how to build a governance stack per jurisdiction. The question becomes how to build an authorization boundary whose artifacts project cleanly into the evidentiary formats each jurisdiction requires, and which jurisdictional projections must be defined.

Four operational consequences follow.

First, investment compounds rather than fragments. Improvements to the authorization boundary improve every projection at once. Improvements to a parallel stack improve only the regime that stack serves. Over multiple regulatory cycles, the projection model produces an architecture that grows in capability while the parallel model produces an architecture that grows in cost.

Second, jurisdictional additions become projection problems rather than architecture problems. A new regime entering scope requires defining its projection mapping. It does not require building a new governance stack. The new projection inherits the properties of the underlying authorization artifact. Onboarding a regime becomes bounded engineering work rather than open-ended program work.

Third, the audit posture improves. Every regime sees the same underlying enforcement reality. Inconsistencies between what was enforced and what was documented stop being a category of risk,

because the documentation is a projection of the enforcement, not a parallel artifact. Regulators in different jurisdictions reviewing the same execution event arrive at consistent reconstructions because they are reading projections of the same authorization decision.

Fourth, adoption shifts from bespoke compliance construction to infrastructure integration. The authorization boundary can be embedded into model-serving systems, agent orchestration frameworks, enterprise AI gateways, workflow engines, and regulated application stacks. Enterprises should not need to rebuild governance for each regime; they should integrate one enforcement layer capable of producing projectable evidence.

The remaining work is real but bounded. Projection mappings per regime must be defined, maintained, and updated as regimes evolve. That work is decoupled from governance architecture. It is evidence-format engineering, not enforcement engineering, and it can be staffed and updated without disturbing the underlying authorization boundary.

Closing

The convergence underneath regulatory fragmentation is not toward better monitoring of AI systems. It is toward provable enforcement at the point where AI becomes real-world action. Once that point is bounded by a non-bypassable authorization layer capable of producing proof-carrying decision artifacts, jurisdictional fragmentation reduces to a question of evidence format rather than governance architecture.

The enterprises that will operate efficiently across the next regulatory cycle are not the ones building the most parallel compliance stacks. They are the ones building one authorization boundary that can project into many.

References

1. Meyman, E. (2026). *On the Impossibility of Observability-Based Authorization* (v1.3.2). Zenodo. https://doi.org/10.5281/zenodo.19647542

2. Meyman, E. (2026). *Observability Is Not Enforcement*. Zenodo. https://doi.org/10.5281/zenodo.18663864

3. Meyman, E. (2026). *From Monitoring to Authorization*. Zenodo. https://doi.org/10.5281/zenodo.18743974

4. Four Tests Standard. https://github.com/edmeyman/4ts-standard

5. Monitoring remains useful for detection, diagnostics, and operational assurance, but it cannot substitute for an authorization verdict emitted before execution.