



A POWDERHOUSE PAPER

Evaluation Precedes Decision

Toward a Persistent Architecture for Institutional Learning

Institutions have built increasingly sophisticated infrastructure for making decisions trustworthy: identity, authorization, governance, audit. That infrastructure treats the decision as the object worth preserving. This paper argues the object is upstream. A decision is the output of an evaluation, and it is the evaluation — the evidence, the analysis, the disagreement, the conditions under which a judgment was formed — that constitutes the durable institutional asset. Preserving decisions produces accountability. Preserving evaluations produces learning.

Chad Kagen
Founder and Chief Product & Innovation Officer, Powderhouse Explorations, Inc.

August 2026 · This paper presents an architectural argument. It is not a product specification.

The Problem

Institutions do not lack information. Global spending on financial market data reached \$49.2 billion in 2025, rising 6.5% on the year. The demand is expanding fastest into private markets — particularly private credit and less liquid assets — where the pressure on providers is to deliver information that is not merely accurate but fully traceable. Data integrity and provenance have become the operative concerns.

The capital confirms it. In 2024 BlackRock acquired Preqin for approximately \$3.2 billion, against roughly \$240 million of recurring revenue — a multiple that prices scarcity in private-market information rather than the revenue itself. PitchBook, the other pole of that market, generated \$583.3 million across the twelve months to June 2024 against a self-estimated addressable market of \$8.8 billion. Institutions are willing to pay substantially for private-market information, and the providers are racing to make it traceable.

Alongside this, a distinct architectural conversation has emerged about the trustworthiness of consequential actions. Watroba (2026) gives it its clearest articulation: as enterprise execution distributes across humans, applications, agents, and machines, authorization should become a dedicated layer between identity and execution, evaluating context, intent, policy, and risk before an action occurs. The paper is careful and, in our reading, correct about the direction of enterprise architecture.

It also contains, almost in passing, the observation this paper is about. Watroba notes that where authorization is embedded in each execution system, compliance can see the result of an action without being able to reconstruct the decision that produced it. He proposes a unified decision record as the remedy: who requested, what was evaluated, why it was approved, what occurred.

That remedy is right, and it is incomplete. A record of what was decided is not a record of how the judgment was formed.

This is not a criticism of that architecture. It is an observation about where the literature currently ends. The adjacent work — zero-trust access, attribute-based access control, AI governance frameworks, shared-signals exchange, and case-based reasoning — converges on making decisions or prior cases trustworthy, retrievable, adaptable, and reconstructible. We are not aware of a framing that treats the preservation of contemporaneous evaluation-time state, together with the conditions under which those evaluations were formed and their subsequent outcomes, as the basis for compounding institutional capability over time. That is the gap this paper addresses, and we would welcome correction on it.

The Inversion

Nearly every architecture in this space treats the decision as the primary object. Authorize, execute, audit, reconstruct, prove. The unit of preservation is the decision and its justification.

But a decision is not a primitive. It is an output. Something is evaluated, and a decision follows. The evaluation contains the evidence considered, the analysis performed, the weightings applied, the conclusions that disagreed with one another, the conditions prevailing at the time, and the uncertainty that was present before the outcome resolved it. The decision is the collapse of all of that into an action.

Evaluations produce decisions. Decisions do not produce evaluations.

The consequence is immediate. A system that preserves decisions preserves the collapse and discards the structure. When the outcome eventually arrives — the company succeeds or fails, the loan performs or defaults, the hire works out or does not — the institution can attribute that outcome to a recorded decision, but not to the evaluation that produced it. It knows what it chose. It cannot recover why the choice looked correct at the time, which conclusions were in tension, or which part of its analysis was wrong.

This is why decision records support accountability but not learning. Accountability asks whether a decision was defensible. Learning asks which part of the reasoning to change. Only the second question requires the evaluation, and only the second question compounds.

The distinction sharpens under time. In private capital, evaluation and outcome are separated by years. A venture assessed today resolves in five to ten. By the time the outcome is observable, the analytical apparatus that produced the original judgment has typically been retrained, upgraded, or replaced, and the conditions under which the judgment was formed are gone. The institution is left comparing an outcome against a reconstruction of its own past reasoning — a reconstruction assembled, unavoidably, by people who now know how the story ended.

Reconstruction after the fact is not the same evidence set, the same model state, the same contextual conditions, or the same disagreement. It is a different object wearing the name of the original.

The Architectural Distinction

The architecture that follows from the inversion has one defining property: it preserves the computational state associated with an evaluation before the outcome is known.

Conventional evaluation systems retain recommendations, outputs, and scores while discarding the computational state that produced them. Systems that learn from outcomes typically do so by retraining: outcomes update model parameters, and in the process the model that produced the original evaluation is overwritten and ceases to exist. Such systems improve by discarding their prior state. They cannot subsequently explain a past recommendation from the state that produced it, re-execute a current model generation against a preserved historical state, or inspect analytical disagreement that was collapsed at inference time.

The disclosed architecture does four things instead. It standardizes heterogeneous evidence into a canonical representation, so that subjects are encoded in a common structure and evaluations remain comparable across time. It preserves the evaluation-time computational state as a persistent object rather than a transient one. It relates that preserved state causally to subsequently observed outcomes, creating a provenance-linked record from evidence through evaluation to result. And it makes those preserved relationships available to future evaluations, permitting replay, calibration, provenance inspection, and institutional learning over states that actually existed.

Because the evaluation-time state is persisted before the outcome exists, later analysis based only on the outcome does not contain the same contemporaneous evidence, model state, conditions, or disagreement.

The claim here is narrow and worth stating precisely. It is not that comparable software cannot be built. A canonical evidence framework, parallel evaluation, an orchestration layer, a modular analytical architecture, an outcome-feedback process — all of these are reproducible. What is not reproducible is the accumulated record: the contemporaneous evidence submitted to a particular system, the actual model and graph state consulted at the time of each evaluation, the independent conclusions and disagreements produced, and the outcome relationships created through operation. Preserved evaluation-time state is not equivalent to history reconstructed after the outcome is known.

What cannot be recreated is the contemporaneous evaluative state from which those judgments arose.

The Consequence

If the inversion holds, the economic question changes shape. The relevant purchase is not analytical software. It is a measurable improvement in the quality of consequential judgment, in a domain where that quality varies enormously.

The variance is documented. Venture capital exhibits the widest performance dispersion of the major alternative asset classes. Across a representative analysis, the median venture fund produced an internal rate of return of 8.2%, while top-quartile funds produced between 17.0% and 79.0%. Most investors in bottom-quartile funds lost capital, at rates between 0.3% and −22.4%. Benchmark aggregations place the top-to-bottom-quartile spread above thirty percentage points in most vintage years — roughly three times the dispersion observed in buyout — which is why manager selection dominates allocation outcomes in this asset class.

Thirty points of net IRR is the distance between institutional excellence and capital destruction, and it is produced by judgment. Not by access to information: the information market is \$49.2 billion and broadly available. The dispersion persists despite the data, which suggests the binding constraint is not what institutions can see but what they can learn from what they have already seen.

This is the argument for preserving evaluations rather than decisions. An institution that preserves only decisions repeats its analytical errors indefinitely, because the errors live in reasoning that was never retained. An institution that preserves evaluations and relates them to outcomes can identify which part of its judgment was systematically wrong, and correct it. The first institution accumulates records. The second accumulates capability.

*Institutions do not compound because they make decisions.
They compound because they preserve, relate, and improve the
evaluations from which those decisions arise.*

Stated that way, the scope of the claim is broader than any single market. Capital allocation is the domain we have chosen to prove it in, because evaluation quality there is measurable, consequential, and separated from outcome by a period long enough to make the reconstruction problem unavoidable. But the structure is not specific to capital. Any institution that evaluates before it acts — and observes results afterward — has the same asset sitting unpreserved.

Conditions for Falsification

An architectural argument that cannot fail is not an argument. The thesis of this paper makes an empirical claim — that preserving and relating evaluations improves subsequent evaluation quality — and that claim is testable. The following conditions are specified in advance of the evidence, and are binding on us.

Outcomes and latency are defined before measurement.

Each measured claim names its outcome event and observation window before measurement begins. Outcomes observed outside a pre-specified window are recorded but do not enter the analysis. No outcome definition is altered after any qualifying event has been observed.

No claim below the evidence threshold.

A minimum cohort of evaluations with observed outcomes is required before any claim of separation is made. Below that threshold the only permitted statement is that evidence is insufficient. Directional language is not a substitute for evidence.

Confounders are named and carried.

Claims are normalized across pre-named confounder classes: stage, sector, vintage and cycle timing, geography, capital access, and founder experience. Every claim carries its normalization method and a statement of residual confounding. A separation that survives only without normalization is not a separation.

The baseline is the incumbent process.

The comparison is not against chance. It is against the institution's own existing evaluation process, with calibration and discrimination reported separately. The system must separate from the practice it proposes to improve.

The cost of evidence is carried in the economics.

Generating decision-grade evidence has a cost — evaluations performed, outcomes tracked, latency carried. That cost is a named line in the operating model, not an overhead absorbed silently.

The claim is suspended if it fails.

If no measurable separation exists by a pre-specified date, at these thresholds and normalizations, the decision-improvement claim is suspended: withdrawn from investor materials, public surfaces, and commercial conversations until separation is demonstrated. It does not remain open pending better data.

The architecture does not ask to be believed. It asks to be checked.

References and Notes

Watroba, M. (2026). The Rise of the Authorization Decision Layer: Why Enterprise Architecture Needs a New Control Point Between Identity and Execution. Industry point of view, publication draft, July 2026.

Burton-Taylor International Consulting (2026). Financial Market Data/Analysis: Global Share & Segment Sizing 2026. Global market data revenue of \$49.2 billion in 2025, growing 6.5%; demand expanding into private markets with increasing emphasis on traceability and provenance.

BlackRock, Inc. (2024). BlackRock to Acquire Preqin, Leading Private Markets Data Solutions Provider. Press release, June 30, 2024. Approximately \$3.2 billion for a business with approximately \$240 million of recurring 2024 revenue.

Morningstar, Inc. (2024). PitchBook segment revenue of \$583.3 million for the twelve months ended June 30, 2024, against a company-estimated addressable market of \$8.8 billion as of February 2024.

CAIS (2026). Performance Dispersion in Alternative Asset Classes. Median venture fund IRR of 8.2%; top-quartile funds between 17.0% and 79.0%; bottom-quartile between 0.3% and −22.4%.

Cambridge Associates benchmark aggregations, as reported 2025–2026. Top-to-bottom-quartile net IRR spread in venture capital exceeding thirty percentage points in most vintage years.

Case-based reasoning literature, generally. Prior cases are represented, indexed, retrieved, adapted to present circumstances, and retained. Cited as the closest adjacent field to the argument advanced here, and distinguished in Section 1: the architecture described concerns preservation of the contemporaneous evaluative object prior to outcome, rather than representation of cases after the fact.

National Institute of Standards and Technology. Zero Trust Architecture (SP 800-207, 2020); Guide to Attribute Based Access Control (SP 800-162, updated 2019); Artificial Intelligence Risk Management Framework 1.0 (2023). Cited as representative of the adjacent literature on policy-based decision infrastructure.

Note on scope. This paper describes an architectural argument and the conditions under which it should be rejected. It does not describe implementation, and it does not characterize the scope of any pending or future patent claim. Subject matter relating to the preservation of evaluation-time computational state and its relation to subsequently observed outcomes is the subject of United States provisional patent applications filed by Powderhouse Explorations, Inc.