

The institutional intelligence gap

Enterprise AI can now reach a financial institution's data. It still cannot apply the definitions, policy, precedent and judgment that decide what the right answer is for that institution.

Access is no longer the bottleneck

Over the past three years financial institutions have connected models to their document stores, data warehouses and collaboration tools. Retrieval works. A model can find the credit memo, the policy manual, the board pack and the spreadsheet behind last quarter's numbers in seconds.

Yet the outputs still need heavy review before anyone acts on them. Senior people rewrite the memo, correct the definition, add the exception the model missed and explain why last year's committee decided differently. The work has moved, but the judgment has not.

The reason is structural. What makes an answer right inside a bank, a fund or an insurer is rarely in any single record. It sits in how the institution defines its terms, which policy applies to which case, what was decided before and why, and who has the authority to depart from the rule.

Five layers, one missing

It helps to describe the modern stack as five layers. The data layer stores and moves records. Models interpret and generate language. Agents chain steps together and call tools. Decisions are the outcomes people and committees are accountable for. Between agents and decisions sits a layer most institutions have never built in machine-usable form: the institution's own knowledge of how it decides.

That layer is not a glossary or a wiki. It is the structured account of concepts, rules, exceptions, evidence requirements, owners and precedent that experienced staff carry in their heads and partly in their documents. When it is missing, every model call improvises it, and every reviewer has to supply it again by hand.

Data layer: where records live and how systems exchange them.

Models: general reasoning and language, trained on the public world.

Agents: sequences of steps, tools and actions.

Institutional knowledge layer: the firm's definitions, policy, precedent and authority, made explicit and governed.

Decisions: the accountable outcomes the institution must defend.

Same data, different right answers

Two lenders can hold identical information about a borrower and reach different, equally defensible conclusions. One treats a covenant breach as a trigger for workout. The other treats the same breach as the start of a pricing conversation, because its relationship strategy, risk appetite and history with that sponsor are different.

A model trained on public data and given access to both lenders' files cannot tell which conclusion belongs to which institution. Nothing in the records says so. The difference is institutional, and it is exactly the difference that matters to the people accountable for the decision.

This is why accuracy against a generic benchmark is a weak test for financial AI. The question is not whether an answer is plausible. It is whether the answer is correct for this firm, under this policy, for this purpose, with evidence the firm accepts.

Where the gap shows up across financial services

The pattern repeats across every part of the sector, though the vocabulary changes.

Commercial lending: policy exceptions, committee precedent and relationship context decide whether a credit is approved, restructured or declined.

Investment banking: house views on comparables, valuation conventions and what a senior banker considers a credible pitch.

Private equity: value-creation playbooks, operating-partner judgment and lessons from earlier portfolio companies.

Insurance: underwriting appetite, referral thresholds and claims precedent that never fit neatly into rules engines.

Asset management: the investment process, research conventions and risk limits that define a house style.

Community banks and credit unions: knowledge concentrated in a few long-tenured people who are approaching retirement.

Why existing approaches fall short

Prompt engineering pushes institutional context into instructions that are long, fragile and invisible to governance. Fine-tuning bakes yesterday's knowledge into model weights that cannot be inspected or changed when policy moves. Retrieval finds passages that look similar to a question, which is not the same as finding the rule that governs it.

Data catalogues and semantic layers help, but they describe tables and metrics. They rarely capture the reasoning around a decision: which exception applies, who can approve a departure, which earlier case is the relevant precedent and why.

Each approach treats institutional knowledge as an input to be stuffed into the model. The alternative is to treat it as infrastructure: owned, versioned, inspectable and available to every person, model and agent that needs it.

What a governed knowledge layer needs

A useful institutional knowledge layer has a small number of properties that distinguish it from documentation.

Firm-specific definitions with scope, owner, jurisdiction and effective dates.

Rules and exceptions linked to the evidence they require and the authority that can override them.

Precedent captured with the fact pattern that made it relevant, not only the outcome.

Provenance on every output, showing which meaning and evidence were applied.

Versioning, so a change in policy can be traced to every workflow it affects.

Permissions, so agents act only within the authority the institution has granted.

Capturing knowledge without stopping the business

The practical objection is effort. Senior experts are the scarcest resource in any financial institution, and nobody will sit through months of interviews to fill a repository.

The workable approach starts from one consequential decision rather than an enterprise taxonomy. It uses the documents, models and past cases the institution already has, asks experts targeted questions only where the record is ambiguous, and validates the result against real cases the experts recognise. The knowledge captured for the first decision becomes the core for the next.

08

A 90-day path

Institutions that make progress tend to follow a similar sequence.

Weeks 1 to 2: choose one decision that is frequent, consequential and contested, such as a credit exception, an investment committee recommendation or an underwriting referral.

Weeks 3 to 6: map the definitions, rules, evidence and precedent behind it from existing material, with short expert reviews.

Weeks 7 to 10: run the decision on historical cases and compare outcomes and reasoning with what the experts actually did.

Weeks 11 to 13: put the governed knowledge behind a live workflow, with human sign-off, provenance and an owner for every rule.

09

The question for leadership

Every financial institution is about to give AI more access. The more important question is what that AI will know about how the institution thinks, and who will own that knowledge.

Firms that make their judgment explicit and governable will be able to scale it across people, models and agents. Firms that leave it implicit will keep paying senior people to correct machines, and will keep losing that judgment every time someone leaves the room.

Bring us one consequential decision.

Penomic makes a financial institution's definitions, policy, precedent and judgment usable by people, models and agents, with provenance on every output.

Request a private briefing at penomic.ai/contact · More research at penomic.ai/insights