

Knowledge archaeology: recovering how the firm actually works

A firm's most valuable financial know-how lives in experts, working files and exceptions—not in a repository waiting to be scanned.

The knowledge is present, but not assembled

An experienced underwriter, investor, controller or claims leader often carries years of pattern recognition that has never been written as a complete method. Parts appear in policy. Parts live in spreadsheets. Parts surface only when a difficult case reaches committee. Some rules are so habitual that experts no longer describe them unless a concrete example forces the distinction into view.

Calling this problem “unstructured data” is too shallow. The material is not waiting to be converted mechanically into triples. Someone must understand why two apparently similar cases were treated differently, which definition prevailed, who had authority to decide and whether the exception became precedent or remained a one-off judgment.

Knowledge archaeology is the human-led practice of recovering that operating logic. It treats interviews, models, policies, emails, working papers and accepted outputs as evidence from which a governed domain model can be constructed and validated.

Begin with competency questions, not a data inventory

A discovery starts with questions the institution must answer reliably: Is this exposure inside the limit? Which product definition applies? Does this covenant adjustment match approved precedent? What evidence is required before this claim moves to the next authority? These competency questions create a boundary around an otherwise unlimited knowledge problem.

Each question reveals the concepts, relationships, calculations, sources and decisions that matter. It also gives the eventual ontology a test. If the model cannot help an authorized workflow answer the question with the expected evidence and caveats, it is not yet operationally useful.

This is materially different from scanning a repository and generating a graph of detected entities. Automated extraction can accelerate candidate identification, but it cannot determine which meaning the business has approved or whether two similar labels represent a synonym, a hierarchy, a local override or a disputed definition.

Follow the artifacts people actually use

Official policy is necessary evidence, but it is rarely the whole operating system. Experts may rely on an Excel model maintained by one team, a committee template, a checklist, a calculation embedded in a presentation, a prior approval paper or a working file that reconciles two systems. These artifacts expose the transformations and exceptions that prose alone often hides.

The archaeologist asks what each field means, where its value comes from, which formulas are considered authoritative and what reviewers change before accepting the result. Version history and annotations can reveal where the business has repeatedly negotiated meaning without ever formalizing the outcome.

The purpose is not to absorb every file. It is to locate the smallest evidence set that explains the decision and to record provenance so the ontology remains connected to the people and artifacts that justify it.

Interview for distinctions and exceptions

Generic interviews produce generic taxonomies. Productive elicitation uses cases. An ontologist asks the expert to compare a routine example with a difficult one, explain what changed, identify the source they trusted and describe the moment when the normal process stopped being sufficient.

Disagreement is useful evidence. If risk, finance and the business use the same term differently, the discovery should not force an artificial consensus. It should represent each definition, its scope and owner, then identify the workflow conditions that select one meaning over another. Some ambiguity belongs in the ontology because it belongs in the institution.

Ask for a normal case, a boundary case and a recent exception.

Trace every material judgment to an owner, source or acknowledged gap.

Separate the written rule from the operating convention and from individual preference.

Record the authority that can resolve a conflict rather than letting the model choose silently.

Turn accepted examples into reusable evaluation cases.

Formalize without losing the business language

The knowledge architecture should express concepts and relationships precisely enough for systems to use while remaining legible to the experts accountable for them. Penomic may use formal semantic techniques internally for identity, hierarchy, constraints and validation, but experts govern the result through clear business definitions, examples and Knowledge Architect—not technical files.

A concept should retain its familiar name, aliases, definition, scope, owner, evidence and review status. A rule should show the conditions that activate it and the exception path when it cannot be applied. A calculation should connect inputs and transformations to the measure it produces. Provenance should survive conversion into machine-readable form.

This dual legibility matters for governance. If only ontology engineers can understand the model, the institution cannot truly own it. If the model remains a workshop diagram, software and agents cannot reliably use it.

Validate through real work

Validation is not a final walkthrough of boxes and arrows. The discovery team runs representative cases through the emerging model. Experts review whether the correct definitions, sources, rules and escalation paths were selected. Failures are classified: missing knowledge, incorrect mapping, ambiguous policy, insufficient evidence or an execution error.

Golden answers are especially valuable when they preserve why an answer was accepted, not only the final text. They become regression tests for retrieval, reasoning and workflow changes. Competency questions can then be evaluated each time the ontology, model provider or agent pipeline changes.

This approach turns expert review into a reusable control. It also distinguishes an ontology problem from a prompt problem, preventing teams from repeatedly tuning language around a missing institutional definition.

A focused first mission should become a working system

The first engagement should not end with a generic transformation report or a technical blueprint alone. For a bounded financial mission, Penomic works with accountable experts to capture the governing definitions, relationships, calculations, evidence, exceptions and authority; connects approved sources and models; tests representative cases; and configures the platform for real work. The usual 4–8 week path depends on mission scope, integrations and controls.

The institution receives a working, reviewable deployment and the knowledge assets behind it: prioritized competency questions, source and owner mappings, accepted examples, evaluation cases and governance decisions. These establish a defensible basis for the recurring platform license and for expanding into adjacent missions.

At scale, the method becomes repeatable. Shared concepts form a semantic core while domain teams own specialist modules. Internal experts, Penomic ontologists and delivery partners can work inside one governance model. The objective is not to centralize every judgment; it is to make institutional meaning discoverable, attributable and usable wherever authorized work occurs.

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