

Your Process Map Shows the Rules. Your Records Show the Exceptions AI Needs

TechRounder PDF Edition

Live article:

<https://www.techrounder.com/business/your-process-map-shows-the-rules-your-records-show-the-exceptions-ai-needs/>

By Vipin PG | Published October 26, 2026 | Updated October 26, 2026 | Format: Article | 6 min read

In brief

Consider a refund process written as three steps: receive the request, check eligibility, and issue the payment. It looks straightforward on a process map.

Consider a refund process written as three steps: receive the request, check eligibility, and issue the payment. It looks straightforward on a process map. Ask the people handling those requests, however, and the exceptions start to emerge.

One customer needs an account review. Another request requires a manager's approval. A third depends on a contract signed years ago, before some of the current employees joined the company.

These situations require employees to look beyond the standard instructions. They have to check the evidence, understand which rules apply, and sometimes ask another person to make the decision. When a business keeps records of those decisions, the resulting history can help explain how its workflows operate in practice.

That history may also be useful when developing or evaluating AI systems that are expected to perform similar tasks. The challenge is finding out whether the records contain enough information to reconstruct what happened and why.

Why workflow exceptions matter for AI

A process document usually explains the expected route through a task. It may say that a support agent can issue a credit when a customer meets certain conditions. The actual case record might reveal that a contract exception changed the decision, or that the employee had to request additional evidence before proceeding.

An AI agent trained to resolve requests quickly could get that situation wrong if its examples only reward fast completion. Sometimes a correct result is an escalation, a refusal, or a request for more information.

Before evaluating an AI system on a business task, it helps to define what competent handling involves. That includes the information available to the employee, the applicable policy, their authority to act, and whether the recorded outcome genuinely resolved the issue.

What useful AI business workflow data contains

A good historical case allows someone reviewing it to follow the decision from the beginning. For a refund or support request, that might mean answering questions such as:

- What did the employee know when the request arrived?
- Which policy or contract version applied at that time?

- Was the account or transaction verified before action was taken?
- Did the case require an approval or escalation, and was that recorded?
- What happened after the decision, and does the outcome record establish that the issue was resolved?

There is a useful parallel in AI agent research. The tau-bench research tests agents working with policies, simulated users, and tools in retail and airline scenarios. Its evaluation compares the database state after an interaction with an expected result. A business archive is different from a controlled benchmark, but the study illustrates why final outcomes and rule-following both matter when assessing an agent.

Not every company will have all this information in one place. A ticketing system may hold the request and resolution, while approval details remain in email or an internal chat. Understanding those gaps is part of assessing what the records can support.

Keep the original decision separate from what happened later

Closing notes often contain facts that were unavailable when the employee first handled a case. Someone reviewing an old complaint may know that a shipment was lost, for example, even though the support agent was still waiting for the carrier's response at the time.

If an AI evaluation supplies that later information at the start, the task becomes easier than the real decision was. The agent appears to have handled a difficult situation correctly because it was given the answer in advance.

For this reason, decision-time information and later outcomes should be stored or labeled separately. Timestamps, policy versions, approval histories, and changes to a case can help reviewers understand the order in which information became available.

What if the records are incomplete?

Older business systems rarely contain a perfect history. Status fields change, attachments disappear during migrations, and some approvals happen outside the system. Duplicate cases and incomplete notes can make it difficult to reconstruct a sequence of events.

These problems don't automatically make the archive unusable. They do limit what can be claimed about it. A collection containing reliable support outcomes but inconsistent approval records may still be useful for some forms of analysis, even if it cannot demonstrate how every exception was authorized.

A practical first step is to inventory the systems involved, the years covered, and the important missing fields. From there, a business can decide whether to improve its internal records, narrow the proposed dataset, or investigate an external licensing assessment.

Check data rights before considering a license

Customer messages, employee communications, contract terms, and internal discussions can all appear in workflow records. Having access to those materials for daily operations does not necessarily give a company permission to license them for external AI training or evaluation.

Any proposed dataset needs an appropriate rights, confidentiality, and privacy review. The business should identify who has authority to approve the use, what its contracts and notices permit, and which records must be excluded. Removing names is often insufficient because other details can still identify individuals.

Licensing terms deserve attention too. Scope of use, exclusivity, duration, sublicensing, retained rights, and deletion obligations can affect how a company may use or license the same material later. These points should be reviewed before signing an agreement or transferring records.

Can a workflow archive have commercial value?

Some companies that work with AI developers are interested in historical business records because they contain examples of how real tasks were handled. These can include customer support histories, sales decisions, engineering activity, and operational workflows.

One business working in this area is SourceX, which says it licenses approved datasets directly from companies for AI training and evaluation. The company handles the agreement and data preparation under the terms of each transaction, with payment terms specified in the signed deal.

SourceX advertises potential payments of up to \$4 million for the largest enterprise datasets. That figure is an advertised upper limit, not a typical payment or an estimate of what a particular archive will earn. Actual eligibility and pricing depend on the scope, history, quality, rights, and commercial terms of the proposed dataset.

According to the company's published eligibility information, its primary focus is businesses that have reached at least 50 full-time employees and have several years of operating history. Its initial fit check asks about the business without requiring records to be uploaded or creating a commitment to proceed.

The company's published licensing information also describes exclusivity conditions for AI training. Any business considering a proposal should check the exact permitted uses, restrictions, payment timing, and ongoing obligations in the agreement rather than relying on headline figures.

Where should a business start?

Choose one workflow in which employees regularly exercise judgment. A refund process, contract approval, support escalation, or similar task can provide a useful starting point. Map the normal route, identify a recurring exception, and locate the records that explain how employees handled it.

Then check whether the archive preserves the information available at each decision, any approvals required, and the eventual outcome. This small review will make missing information easier to spot and provide a more realistic picture of the archive's usefulness.

For businesses with several years of connected operational records, SourceX offers a no-upload company fit check for exploring whether a licensing assessment is appropriate. A sensible starting point is an internal inventory and approval from the people responsible for the records.

About the founder

Noah Loul is the founder and CEO of SourceX. He previously founded B2B Rocket, which was acquired by Blackpearl Group. His current company focuses on sourcing business datasets for AI training and evaluation through agreements with the organizations holding the records.

References

1. arxiv.org - abs / 2406.12045 - <https://arxiv.org/abs/2406.12045>
2. sourcex.si - https://sourcex.si/?utm_source=techrounder.com&utm_medium=sponsored_article&utm_campaign=sourcex_publication_oct2026&utm_content=24_ai_business_workflow_data