

WHERE TO START WITH AI · NO. 2

Reviewing expense reports

An expense review assistant, rolled out in two stages: first it prepares a recommendation for the reviewer, then it approves compliant reports below a set threshold on its own.

Stage 1 · the reviewer decides, with AI support

Stage 2 · AI approves compliant reports under a threshold



WHY START HERE

It meets all four tests of a good first workflow

✓ High volume

Hundreds of reports a month across about ten categories.

✓ Clear owner

The CFO's organization, which already sets the expense policy.

✓ Clear output

A reasoned recommendation: approve, correct, or reject.

✓ Known pain point

For reviewers and employees alike.

Written rules. The assistant checks each report against an expense policy the company already has. The risk does not go away. It becomes more visible, measurable, and manageable once the rules, the scope, and the escalation thresholds are defined.



ON ONE CONDITION

Low risk, but only within a strict scope

- 1 A written expense policy**
approved, current, and consistent.
- 2 Amounts below a set threshold**
above it, the reviewer always decides.
- 3 No sanctions or judgments**
on employees left to AI.
- 4 Unusual, sensitive, or ambiguous cases**
routed to finance.

Outside this scope, it becomes a different use case with a different risk profile.



THE WORKFLOW AS IT RUNS TODAY

Six steps, all checked by hand

Example Industries Inc. · 600 employees · fictitious company

1	Report submitted by the employee with photographed receipts	
2	Receipt check legibility, date, amount, VAT	● Illegible or missing receipts
3	Policy check limits, categories, guests	● Rules reread for every report
4	Questions to the employee when a receipt is missing	● Back-and-forth by email
5	Manager approval then finance approval	● Double approval every time
6	Posting and reimbursement export to accounting	● Rejection reasons not recorded

≈ 450
expense reports a month

≈ 80%
compliant as submitted

≈ 6 days
from submission to reimbursement

● friction point · illustrative figures

Illustrative case: fictitious company; all figures are approximate.



BEFORE YOU START

Get the policy in order

The work starts by clarifying the expense policy: limits by category and country, required receipts, exceptions, and approval thresholds. Each rule has an owner and a version date. The tool comes after.

1

Limits by category and country

2

Required receipts

3

One owner per rule

4

Dated versions

The tool comes after.



STAGE 1 · HUMAN IN THE LOOP

The assistant recommends, the reviewer decides

The report comes in
through the same tool as today.



The assistant prepares a recommendation
reading the receipts and citing the rule applied.



The reviewer decides
approves, corrects, or asks for a receipt.



Every correction is logged
category by category.

What this stage produces: a category-by-category measure of how accurate the recommendations are.

Estimated duration: 2 to 3 monthly cycles.



STAGE 2 · COMPLIANT REPORTS BELOW THE THRESHOLD

It approves compliant reports below the threshold

THE ASSISTANT APPROVES WHEN

- The expense category is allowed
- Receipts are legible and consistent
- The amount is under the limit and threshold

IT ROUTES TO FINANCE WHEN

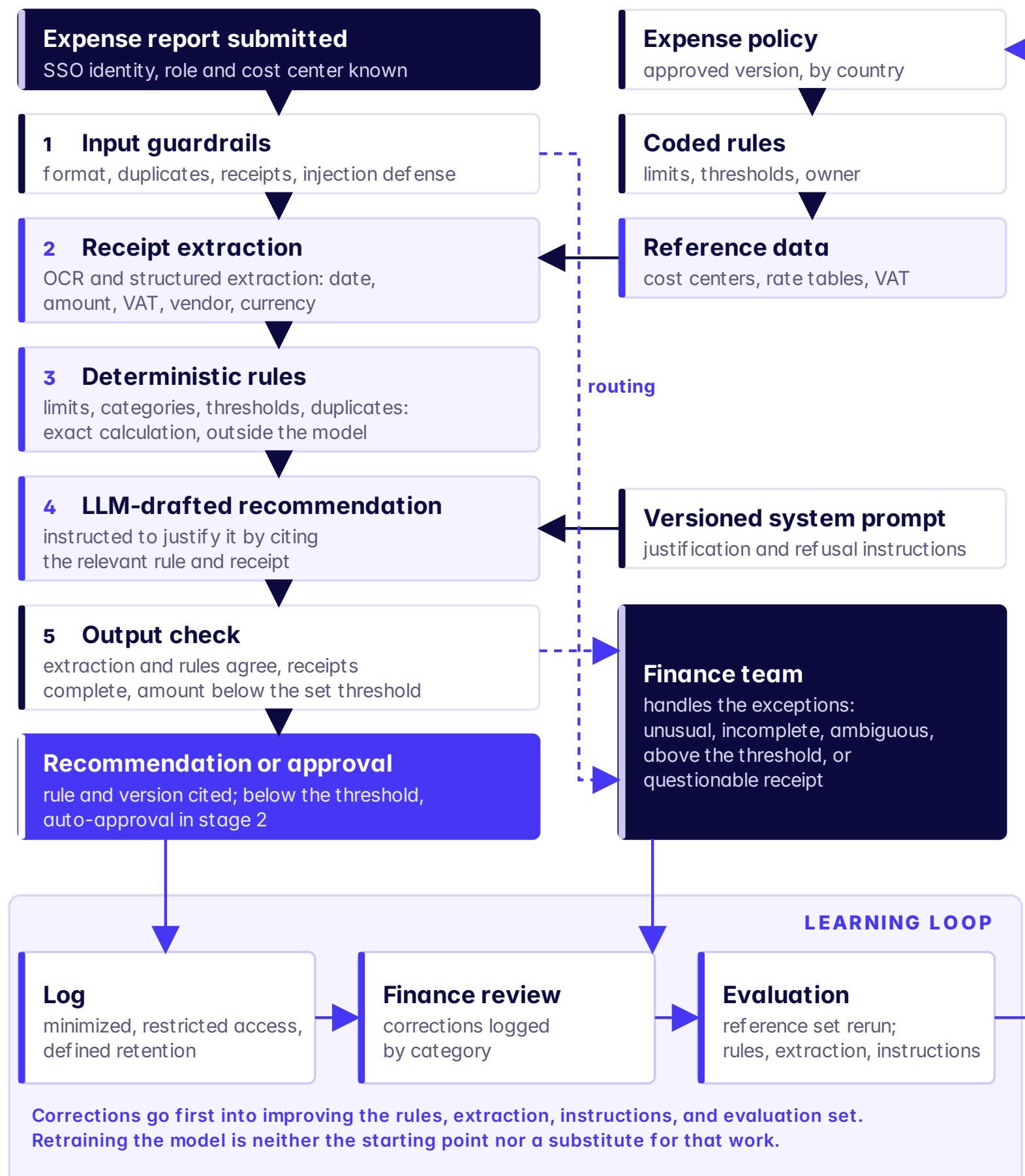
- The report is unusual, incomplete, or ambiguous
- A receipt is missing, inconsistent, or suspect
- The expense involves entertainment, outside guests, travel abroad, a large sum, or similar

Every week, finance reviews a sample of approved reports and keeps the policy up to date.



UNDER THE HOOD · EXTRACTION, RULES, LLM

Reference architecture: how the system stays under control



UNDER THE HOOD · CONTROLS AND SECURITY

The controls required before any autonomous answer

1 Evaluation

A reference set of about 200 past reports, pseudonymized or anonymized as the test requires, each with the expected decision. Metrics: extraction accuracy, recommendation accuracy, correct rule cited, and the rate at which exceptions are detected and correctly routed. Rerun whenever the model, a rule, or the instructions change.

2 Observability

Every report processed is logged: extracted fields, rules applied, recommendation or approval, routing decision. Receipts and personal data: log minimized, access restricted to authorized roles, retention defined by purpose and applicable obligations.

3 Identity and access

SSO (SAML or OIDC). Segregation of duties: whoever submits a report can neither approve it nor change the rules. Access rights enforced from receipt intake onward.

4 Data

Hosting and data transfers verified for the chosen solution; contractual terms barring use of the data for model training; encryption; receipt retention in line with accounting law; documented Swiss FADP assessment.

5 Application security

Protection against prompt injection, including in receipt text. Scope limited to three outcomes: recommend, approve within the authorized scope, or route to finance. Accounting alone triggers payment.

Three options: the module in your existing expense tool, a managed document-extraction platform, or a custom build (extraction, rules engine, model API). The choice comes after the policy audit, not before.



MOVING FROM STAGE 1 TO STAGE 2

A decision based on evidence, category by category

The move to stage 2 should not be decided in advance. It depends on what finance observes during stage 1: the accuracy of the recommendations, the corrections needed, and the system's ability to recognize reports that fall outside its scope.

WHAT FINANCE OBSERVES

- Accuracy of extraction and recommendations
- Corrections needed, by category
- Exceptions correctly routed
- No significant incident within the tested scope
- A current policy, each rule with an owner

Travel expenses could move to stage 2 while business meals stay in stage 1.



THE GOAL, AND HOW TO MEASURE IT

The goal is not to take review away from the finance team. It is to cut reimbursement time for compliant reports, stop the same checks from being redone by hand, and focus reviewers' time on expenses that call for context and judgment.

Time to reimburse a compliant report

≈ 6 days → **≈ 2 days**
today (illustrative) illustrative stage 2 target

Also tracked: the share of reports approved without finance involvement,
40–50% (illustrative target)

Data protection

Swiss FADP: purposes, data processed, employee notice, hosting, and transfers clarified before launch.

Governance

The CFO's organization owns the outcome and reviews a sample of approved reports every week.



In this example, AI does not take review away from the finance team.

It steps in first where a written rule and a legible receipt are enough to decide.

So the useful question is not “Which tool should we buy?”

It is “Which repetitive work can we have AI assist with, measure, and then delegate under clear controls?”

What repetitive work in your company already meets these conditions?