WHERE TO START WITH AI · NO. 5

Matching supplier invoices

An invoice matching assistant, rolled out in two stages: first it flags discrepancies for the accountant, then it posts invoices that match or fall within tolerance, below a set amount, on its own.

Stage 1 · the accountant decides, with AI support

Stage 2 · AI posts qualifying invoices, AP handles exceptions

Illustrative case: fictitious company; all figures are approximate.

WHY START HERE

It meets all four tests of a good first workflow

High volume

About 1,800 invoices a month from some 300 suppliers.

Clear owner

The CFO's organization, which already runs accounts payable.

Clear output

A match summary: matched, discrepancy, or needs review.

Known pain point

For accounting and suppliers alike.

Records that already exist. The assistant compares each invoice with the purchase order and goods receipt the company already records. The risk does not go away. It becomes more visible, measurable, and manageable once tolerances, scope, and escalation thresholds are defined.

Illustrative case: fictitious company; all figures are approximate.

ON ONE CONDITION

Low risk, but only within a strict scope

1. **POs and goods receipts in the ERP**
   complete and up to date.

2. **Amounts below a set threshold**
   above it, the accountant always decides.

3. **No payment or bank account change**
   triggered or made by AI.

4. **Out-of-tolerance, disputed, or ambiguous**
   invoices routed to accounts payable.

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

Illustrative case: fictitious company; all figures are approximate.

THE WORKFLOW AS IT RUNS TODAY

Six steps, all done by hand

Example Distribution Inc. · 450 employees · fictitious company

1	**Invoice received from supplier** PDF by email or portal	
2	**Invoice entry** supplier, amounts, VAT	● Lines rekeyed by hand
3	**PO lookup** and goods receipt	● Records in separate systems
4	**Line-by-line comparison** quantities, prices, terms	● Tolerances applied ad hoc
5	**Questions to buyers** or the warehouse on discrepancies	● Back-and-forth by email
6	**Posting** then payment run	● Discrepancy reasons not recorded

≈ 1,800
invoices a month

≈ 75%
match with no discrepancy

≈ 9 days
from receipt to posting

● friction point · illustrative figures

Illustrative case: fictitious company; all figures are approximate.

BEFORE YOU START

Get the data in order

The work starts by getting the data in order: POs and goods receipts entered on time, duplicate supplier records cleaned up, discrepancy tolerances written down by purchase category, and exceptions defined. Each rule has an owner and a version date. The tool comes after.

1
Current POs and receipts

2
Written tolerances

3
One owner per rule

4
Dated versions

> **The tool comes after.**

STAGE 1 · HUMAN IN THE LOOP

The assistant flags issues, the accountant decides

The invoice comes in
through the same process as today.

↓

The assistant prepares a match summary
extracting the lines and citing each discrepancy.

↓

The accountant decides
approves, corrects, or checks with the buyer.

↓

Every correction is logged
by purchase category.

What this stage produces: a measure, by purchase category, of how accurate the proposed matches are.

Estimated duration: 2 to 3 monthly cycles.

Illustrative case: fictitious company; all figures are approximate.

STAGE 2 · QUALIFYING INVOICES BELOW THE THRESHOLD

It posts qualifying invoices, accounts payable handles exceptions

THE ASSISTANT POSTS WHEN

- Invoice, PO, and goods receipt agree
- Any discrepancy is within written tolerances
- The amount is below the threshold and the supplier is on file

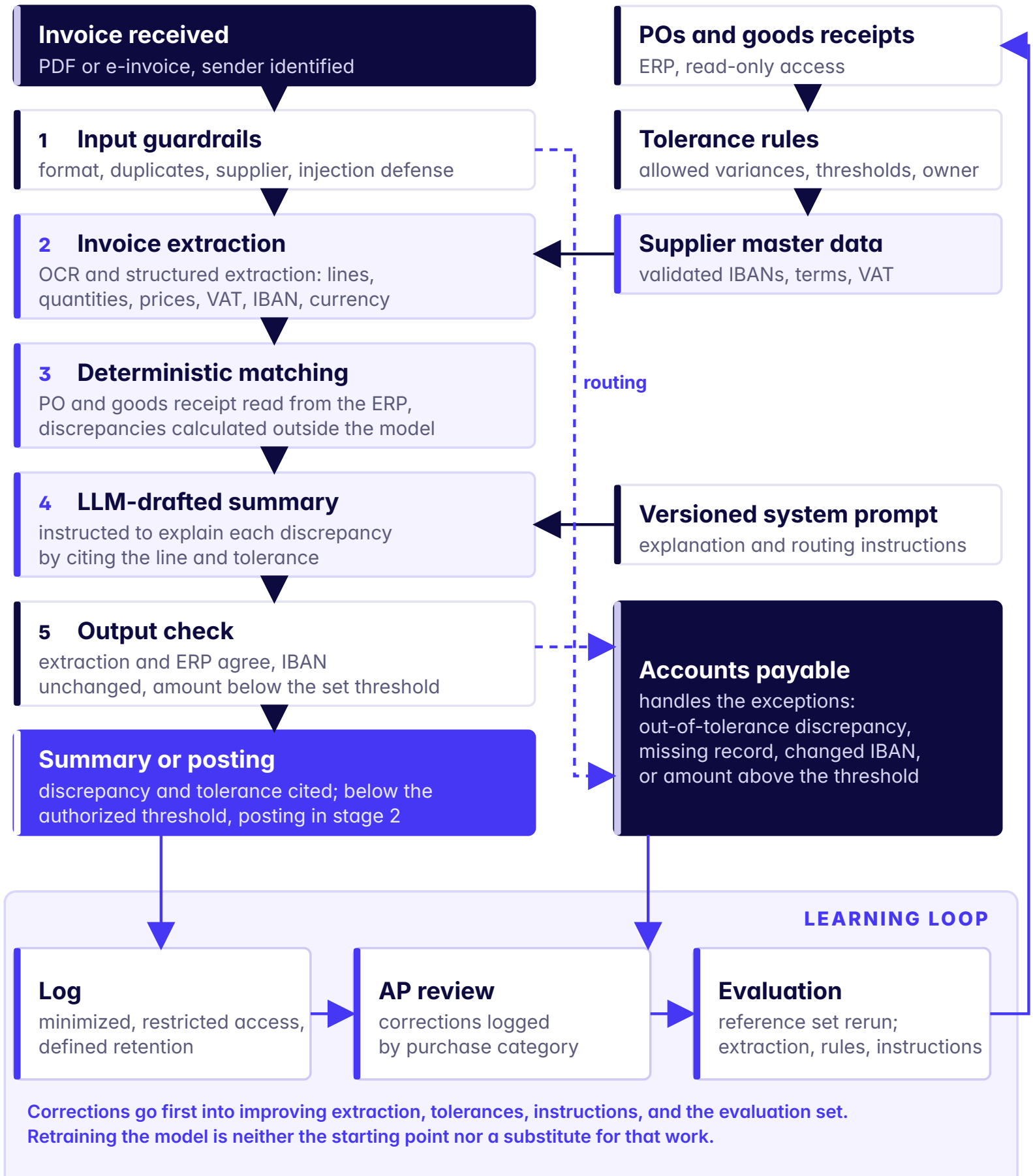
IT ROUTES TO ACCOUNTS PAYABLE WHEN

- The invoice is unusual, incomplete, or ambiguous
- A discrepancy exceeds tolerance or a record is missing
- The invoice involves a new supplier, changed bank account, a credit memo, a large amount, or another red flag

> Every week, accounts payable reviews a sample of posted invoices and keeps the tolerances up to date.

UNDER THE HOOD · EXTRACTION, MATCHING, LLM

Reference architecture: how the system stays under control



Illustrative case: fictitious company; all figures are approximate.

UNDER THE HOOD · CONTROLS AND SECURITY

The controls required before any autonomous answer

1 **Evaluation**
A reference set of, for example, 300 past invoices, minimized as needed, each with the expected match. Metrics: extraction accuracy, discrepancy accuracy, correct tolerance 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 invoice is logged: extracted fields, calculated discrepancies, summary or posting, routing. Supplier and bank 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 places an order cannot approve the invoice; tolerances and supplier records are managed by separate roles. ERP read access; write access limited to authorized postings below the threshold.

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

5 **Application security**
Measures against prompt injection, including in invoice text. Scope limited to three outcomes: propose a match summary, post within the authorized scope, or route to AP. Only treasury releases payments.

> Three options: the matching module in your existing ERP, a managed document-extraction platform, or a custom build (extraction, rules engine, model API). The choice comes after the data audit.

Illustrative case: fictitious company; all figures are approximate.

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 accounts payable observes during stage 1: the accuracy of the proposed matches, the corrections needed, and the system's ability to recognize invoices that fall outside its scope.

WHAT ACCOUNTS PAYABLE OBSERVES

- Accuracy of extraction and discrepancies
- Corrections needed, by category
- Exceptions correctly routed
- No significant incident within the tested scope
- Current tolerances, each with an owner

> Office supplies could move to stage 2 while professional services stay in stage 1.

Illustrative case: fictitious company; all figures are approximate.

THE GOAL, AND HOW TO MEASURE IT

The goal is not to take control away from the accounting team. It is to cut processing time for clean-match invoices, stop the same comparisons from being redone by hand, and focus the team's time on discrepancies that call for context and judgment.

Time to post a clean-match invoice

≈ 9 days → **≈ 2 days**

today (illustrative) — illustrative stage 2 target

Also tracked: the share of invoices posted without manual intervention, **50–60%** (illustrative target)

Data protection	Swiss FADP: purposes, data processed, notice to affected individuals, hosting, and transfers clarified before launch.
Governance	The CFO's organization owns the outcome and reviews a sample of posted invoices every week.

Illustrative case: fictitious company; all figures are approximate.

In this example, AI does not take control away from the accounting team.

It steps in first where the PO, the goods receipt, and the invoice match or stay within written tolerances.

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?**

Sergio Castagna Consulting.
sergiocastagna.com · info@sergiocastagna.com · Geneva

Sergio Castagna Consulting is a member of the OpenAI Partner Network and a registered partner in the Google Cloud Partner Network. The tool comes after the problem is defined.