

AI and automation · Workflow decisions

Decide how much a workflow step should do on its own

Take one step in a workflow the business already runs, most of it already automated, where someone has proposed letting an AI agent act. Decide what a rule can settle, what an assistant can propose, what a person still checks or decides, how exceptions come back, and which old work stops. Finish with one bounded decision, not an upgrade.

What you finish with: One workflow step with the task defined, the volumes counted on one unit of work, the rule that settles part of it, the assistance that proposes the rest, the human check that stays, the exception path, the work that stops, and the defined subset, evidence and authority under which the step may be allowed to do more.

[Read this guide online](#)

WORKFLOW / AUTHORITY

Separate assistance from permission.

- | | | |
|----|---|-----------|
| 01 | Rule settles eligible items
Every condition met; no other hold. | RULE |
| 02 | Assistant proposes the rest
Records, reason and draft action. | PROPOSAL |
| 03 | Person checks and commits
Approve, correct or return the exception. | AUTHORITY |

Allow more automatic action only for a defined subset, with evidence, limits, a route back and the owner's decision.

01 Start with the task and the counts, not the technology

Name the step by the task it performs and the decision it ends in: matching an invoice to a purchase order, releasing a held order, classifying a support request. Count it: items a month, the share the current automation handles cleanly, the share that lands with a person, and the minutes a person spends on each. Take the counts from the last three months of records, not from the proposal, and count in one unit throughout: if the person's task is an invoice, count invoices, and count an invoice as settled only when every hold on it is settled.

A step that lands 40 items a month with a person is a different decision from one that lands 400, and the proposal to add an agent should be judged against the counts, not against the idea.

02 Settle with a rule what a rule can settle

Look at the items that land with a person and ask how many could be resolved by a written rule the business is willing to stand behind: a tolerance, a threshold, a list. Where the business can define a stable rule, use that rule to settle the eligible cases and inspect the exceptions. Write the rule as a business definition with its owner: what it allows, what it never allows, and who changes it. Then count the items that pass the whole rule, every condition and exclusion, and have no other hold; only those items leave the person's queue. The items the rule cannot settle are the ones worth considering for assistance.

03 Decide what an assistant proposes and what a person still decides

For the remaining items, set out the three parts plainly. What the system does: gather the related records, propose the resolution and the reason, and draft the action. What a person checks or decides: the person reads the proposal and the reason, approves or corrects it, and remains the one who commits a posting, a payment, a release or an external message. How exceptions return: any item the system cannot resolve with the records it has comes back to the person with the reason, and any item touching a third party goes to the owner of that relationship.

Which old work stops: name it, because if the person keeps doing the manual lookup beside the proposal, the step costs more than before. Letting the system act without a person is a later decision, taken on evidence, by the person who holds the authority for that action, and only for a defined subset with its own limits, records, exceptions and route back.

04 Test it with counts, then decide on evidence whether the step may do more

Run the step for a stated period with the rule and the assistance in place; the period is the business's choice for that step, long enough to see a full cycle of the work. Count items settled by the rule, proposals accepted unchanged, proposals corrected, items returned as exceptions, and the total time on the new route measured the same way as the old, including review, corrections, rework and the report. An accepted-unchanged rate says what people accepted; it does not say the proposals were right, because people can accept a wrong proposal.

So define in advance the subset the system might act on alone, with the action permitted, the limit per item, the records that must be present, the exceptions that still return, and the route back when an automatic action is wrong: stop the subset the same day, reverse and correct the item, and let the authorized owner decide when it restarts. Have someone independent of the proposal check a sample drawn from that subset against the underlying records, recording how the sample was chosen and which errors were looked for. A clean sample is evidence about the sampled items; it does not prove the unsampled ones.

The person who holds the authority for the action decides whether the evidence and the remaining risk justify the bounded change; insufficient evidence keeps the action with a person. If the evidence does not support automatic action, retain the person's approval. Keep the assistance only if measured time, quality and operating cost justify it; otherwise adjust it or return to the existing route.

Example decision

Invoice mismatches in accounts payable

A company processes about 3,000 supplier invoices a month through an existing matching bot. Twelve percent, 360 a month, fail the match and land with an accounts payable clerk, who spends about 15 minutes on each: 360×15 is 5,400 minutes, 90 hours a month. A proposal on the table is an AI agent that reads each mismatch, checks the receipt and the price list, and posts the correction.

The rule:

Counted invoice by invoice from the last three months, 140 of the 360 a month have a price difference under 25 dollars and under 2 percent of the line as their only hold: the supplier has at least twelve months of history, no line on the invoice is over 5,000 dollars, the receipt is on file and quantities match.

The controller writes the tolerance rule: an invoice meeting every one of those conditions is approved at the invoiced price and reported weekly; the rule never applies to a supplier with under twelve

months of history, to any invoice with a line over 5,000 dollars, or to any invoice carrying another hold; the controller owns the limits.

Work that stops outright: the clerk's lookup on those 140 invoices, 140×15 is 2,100 minutes, 35 hours a month, a task-time estimate before the weekly report's review time is counted.

The assistance:

For the remaining 220, the system gathers the purchase order, the receipt and the price list, proposes the resolution with its reason, and drafts the correction. The clerk reads the proposal, approves or corrects it, and posts; the clerk's posting authority does not move.

Estimated clerk time falls from 15 minutes to 5 per invoice: 220×5 is 1,100 minutes, 18.3 hours, against 220×15 is 3,300 minutes, 55 hours, an estimated difference of 36.7 hours a month of task time, to be measured, not assumed, and before review and correction time.

Exceptions: an invoice with no receipt on file returns to the clerk with the reason; a disputed price returns to the buyer who owns the supplier; nothing is posted without a person.

Work that stops:

The clerk's manual lookup for rule-settled invoices; the weekly mismatch spreadsheet, replaced by the system's weekly report of rule approvals and proposals.

Test, evidence and authority:

Two months, the company's choice for this step. Counts kept by the clerk: rule-settled invoices, proposals accepted unchanged, proposals corrected, exceptions returned, and total minutes on the route including review of the weekly report and any corrections.

The subset the system might post alone is defined now: price corrections under 100 dollars per invoice, on suppliers with twelve months of history, with no line over 5,000 dollars, receipt on file, quantities matched and no other hold; the action permitted is posting the correction into the ordinary payment run, never a payment outside it; anything missing a record or touching a dispute still returns to a person.

The controller, who holds posting authority, decides at the end of the two months on the evidence of a sample drawn from proposals in that subset, chosen as every tenth item plus every item between 50 and 100 dollars, checked by the controller's own team against the purchase order, receipt and price list for wrong price source, wrong receipt, wrong supplier and wrong period.

The accepted-unchanged rate is reported beside it as an indicator, not as the proof, because the wider assisted population does not authorize this subset. A clean sample supports the change for the sampled class and says nothing about invoices outside it.

If the controller allows it, the clerk reviews the daily report of automatic postings and the continuing sample rather than every invoice; if the sample or a supplier query finds an incorrect automatic posting,

automatic posting stops the same day, the clerk reverses and corrects it, and the controller decides when it restarts.

If the evidence does not support it, posting stays with the clerk, and the assistance itself continues only if the measured total time, quality and running cost justify it; otherwise it is adjusted or the old route returns. The accounts payable manager owns the step; the controller owns the rule, the subset and the sample.

Your working record

WORKFLOW STEP DECISION

Field	Your record
Step, by task and the decision it ends in:	
Counts from the last three months, in one unit (items a month, handled cleanly, landing with a person, minutes each):	
Rule that settles part of it (what it allows, what it never allows, owner):	
Items that pass the whole rule with no other hold, and the work that stops:	
Remaining items, what the system gathers and proposes, what the person checks and commits:	
Exceptions and where they return:	
Old work that stops beside the assistance:	
Test period, counts kept, and total time on both routes:	
Subset the system might act on alone (action, limit, records required, exceptions, route back when wrong):	
Evidence for that subset (sample, how chosen, errors looked for, who checks):	
Who holds the authority to allow it:	
If the evidence does not support it, what stays and what is reviewed:	

[Open the guide to copy this record into your team's document](#)

Related reading

[The AI Use Case Dashboard](#)

For the items an assistant proposes, the dashboard's cases show a prompt, a sample and the checks a person runs before anything is posted.