

The AI Use-Case Shortlist

A scoring worksheet for deciding which processes to automate first — before you spend anything finding out.

Most failed AI projects were not badly built. They were badly chosen. The team picked the process that was most annoying, or most visible to leadership, rather than the one most likely to work — and discovered the mismatch nine months and a six-figure budget later.

Choosing well is a far cheaper problem to solve than building well. This worksheet ranks candidate processes against the six factors that actually predict whether an automation effort succeeds, so the conversation stops being about enthusiasm and starts being about evidence.

How to use it. Pick three to five candidate processes. Score each separately against all six sections, total the scores, then read the verdict bands. Do not score a process you cannot describe end to end in five minutes — that inability is itself the finding.

Before you score: name the process properly

"Invoice processing" is not a process — it is a department. "When a supplier invoice arrives by email, someone extracts the line items, matches them to a purchase order, and either approves it or routes the exception" is a process. You can score that.

If your candidate does not survive being written as trigger, action and output, split it. Most "one big process" candidates are three or four smaller ones with very different automation prospects, and the good one gets dragged down when they are assessed together.

Candidate process	Your answer
Trigger — what starts it	
Action — what happens	
Output — what comes out	
Who does it today	
How often it runs	

The six-factor score

Eighteen statements across six factors, scored 0 to 2. Maximum 36. Be honest rather than optimistic — a low score is a useful result, not a failed test.

1. Volume and repetition

0 = not true · 1 = partly true · 2 = clearly true	Score
This process runs at least weekly, and usually more often.	0 1 2
The same steps are followed each time, with few genuine exceptions.	0 1 2
More than one person spends meaningful time on it every week.	0 1 2
Section subtotal (out of 6)	_____

2. Rules over judgement

0 = not true · 1 = partly true · 2 = clearly true	Score
A competent new hire could be taught the decision rules in under a day.	0 1 2
Where judgement is needed, it sits at a small number of identifiable points.	0 1 2
Someone could explain why any given past decision was made.	0 1 2
Section subtotal (out of 6)	_____

3. Input quality

0 = not true · 1 = partly true · 2 = clearly true	Score
Inputs arrive in a consistent, machine-readable form (not scanned or freeform).	0 1 2
The data needed already lives in a system, not in someone's head or inbox.	0 1 2
You could produce twenty real past examples of inputs and correct outputs today.	0 1 2
Section subtotal (out of 6)	_____

4. Cost of being wrong

0 = not true · 1 = partly true · 2 = clearly true	Score
A wrong output is caught before it reaches a customer or a ledger.	0 1 2
A wrong output is reversible without legal, financial or safety consequences.	0 1 2
There is a human review step you would keep for at least the first six months.	0 1 2
Section subtotal (out of 6)	_____

5. Measurable outcome

0 = not true · 1 = partly true · 2 = clearly true	Score
You can state today's baseline as a number: hours, cost, cycle time or error rate.	0 1 2
You would know within one month whether the change worked.	0 1 2
Someone owns that number and cares whether it moves.	0 1 2
Section subtotal (out of 6)	_____

6. Organisational readiness

0 = not true · 1 = partly true · 2 = clearly true	Score
The team doing the work today has been told this is being explored.	0 1 2
A named person can approve changes to how this process runs.	0 1 2
No active migration or reorganisation is about to reshape this process.	0 1 2
Section subtotal (out of 6)	_____

Totalling and reading the result

Factor	Subtotal /6
Volume and repetition	_____
Rules over judgement	_____
Input quality	_____
Cost of being wrong	_____
Measurable outcome	_____
Organisational readiness	_____
TOTAL (out of 36)	_____

The lowest section matters more than the total

A process scoring 30 overall with a 1 in Input Quality is not a strong candidate — it is a candidate with one fatal problem. Automation efforts fail at their weakest factor, not their average. Read your lowest subtotal first and treat it as the thing to fix, or the reason to walk away.

Verdict bands

28–36 — Strong candidate

Well suited to automation, provided no single factor scored below 3. Move to a scoped assessment: confirm the data exists in the form you believe it does, agree the baseline number, and define what human review looks like for the first six months.

20–27 — Viable with preparation

Usually worth doing, but not yet. Fix your two lowest factors first. In practice this is nearly always input quality or organisational readiness, and both are cheaper to address before a build than during one.

12–19 — Not yet

Something structural is in the way. Re-score once it is resolved. Automating a process this shape tends to produce a system that works in the demo and quietly stops being used within a quarter.

Below 12 — Wrong candidate

Pick a different process. This is the most valuable result the worksheet can give you, and by far the cheapest place to receive it.

A worked example

A 60-person distribution business scored three candidates. The figures are illustrative, but the shape of the result is the one that recurs most often.

Factor	Quote approvals	Supplier invoices	Customer emails
Volume and repetition	3	6	6
Rules over judgement	2	5	2
Input quality	4	5	1
Cost of being wrong	1	5	3
Measurable outcome	3	6	2
Organisational readiness	4	4	5
Total	17	31	19

Quote approvals (17) — not yet. The lowest factor was cost of being wrong: an incorrectly approved quote goes straight to a customer as a commitment. That single score disqualifies it as a first project, however appealing the time saving looked.

Customer emails (19) — wrong candidate, despite the volume. This was the process leadership kept raising, and it scored top marks on volume. But inputs were freeform text and the rules were genuinely judgement-heavy. High enthusiasm, low suitability.

Supplier invoices (31) — clear winner. Not the process anyone had complained about. Structured inputs, teachable rules, an internal-only failure mode, and a baseline the finance lead already tracked. It shipped, the number moved, and the credibility it earned made the harder projects fundable.

The pattern to notice: the process that won was the one nobody had nominated. If your scoring produces the answer everyone already assumed, check whether you scored honestly or simply justified a decision that had already been made.

Four objections you will hear

"We should start with the biggest problem."

Biggest problems are usually big because they are hard — high judgement, high consequence, tangled inputs. Starting there makes your first attempt your hardest attempt, run by a team with no experience of what these projects need.

"The vendor demo handled something just like this."

Demos run on clean data the vendor prepared. Your Input Quality score is the honest version of that question. If it is low, the demo told you nothing about your situation.

"We don't have time to score things, we need to move."

Scoring three candidates takes an afternoon. Choosing wrong costs a quarter at minimum, and usually costs the organisation its appetite for trying again.

"Can't we just automate the whole department?"

Departments are not processes. Every successful programme here is a sequence of narrow wins that compound. Breadth comes later, from having shipped something that worked.

What the first ninety days should look like

1. **Weeks 1–2 — Validate the inputs.** Pull twenty real historical examples with their correct outputs. If assembling those takes more than a couple of days, your Input Quality score was optimistic and the project should pause here rather than later.
2. **Weeks 2–3 — Fix the baseline.** Write down today's number and who owns it. Projects without an agreed baseline cannot be judged, which means they cannot be stopped, which is how they become permanent.
3. **Weeks 3–8 — Build narrow.** Handle the common path only and route every exception to a person. Resisting edge cases in version one is the single biggest predictor of shipping.
4. **Weeks 8–12 — Run in parallel with review.** Keep the human check and measure disagreement between system and reviewer. That gap is your real accuracy signal, and far more informative than any test-set figure.
5. **End of quarter — Decide honestly.** Did the baseline number move? Expand, adjust, or stop. All three are legitimate; only drifting is not.

Keep the review step longer than feels necessary. The most common expensive mistake is removing human oversight the moment accuracy looks acceptable in testing. Live inputs drift in ways test sets do not, and the resulting failures are usually discovered by a customer rather than by you.

Next step: the AI Readiness Assessment

This worksheet tells you which process to pursue. The Atomix AI Readiness Assessment (\$997) takes your top-ranked candidate and validates it properly — data availability, integration surface, review design, baseline definition and a costed implementation path. The difference between a shortlist and a plan.

General business guidance only. This worksheet is a planning aid, not professional, legal, financial or regulatory advice, and it does not assess compliance with any AI-specific regulation or standard. Automation decisions involving personal data, employment outcomes, credit, health, safety or other regulated activity carry obligations this scoring does not evaluate — take qualified advice on your own circumstances before proceeding.