

90 DAYS TO YOUR FIRST AI EMPLOYEE

The quarter-long plan for putting one governed AI agent to work — without betting the business on it

Companies stall on AI for a year debating strategy, then panic-buy something broad and regret it. Here's the alternative: **one AI agent, one bounded job, one quarter** — run like the hiring process it actually is. Small enough to be safe. Real enough to teach you everything the think-pieces can't.

WEEKS 1-2 — PICK THE JOB (THIS DECIDES EVERYTHING)

You're not looking for the most impressive use of AI. You're looking for the best *first hire*. Score your candidate tasks against five tests:

TEST	YOU WANT
Boring	Work your team is glad to shed — triage, reconciliation, chasing, assembling
Bounded	Describable in two sentences, touching 1–2 systems
High-volume	Happens daily/weekly, so one quarter generates real evidence
Checkable	A human can verify an output in under a minute
Forgiving	A caught mistake costs an correction, not a client

Classic winners: intake triage and drafting, matching orders/invoices/receipts and flagging exceptions, deadline and renewal watching, assembling recurring reports. Classic losers for a first hire: anything customer-facing without review, anything touching money without approval, "just automate everything about X."

Deliverable: one paragraph — the job description, in business language. If leadership can't read it, it's not ready.

WEEKS 3-6 — SET THE TERMS OF EMPLOYMENT

Before the agent does anything, write its one-page handbook (scope, access, approvals, logging, off switch). The non-negotiables:

- Badge, not master key:** it gets the permissions of the most junior person who could do this job. Enforced by the system, not by promises.
- Approval matrix:** anything leaving the company, moving money, or changing a record that matters pauses for a named human. Start strict — loosening later is a decision; tightening later is an incident report.
- Logbook from day one:** every action, timestamped, reviewable. No log, no launch.
- Data terms in writing:** your data isn't training anyone's model. Enterprise terms, not a sales email.
- Test the off switch** before go-live. Flip it. Watch work stop. Now it's real.

WEEKS 7-10 — THE SUPERVISED PILOT

Run the agent on real work with **100% human review** — every output checked before it counts. Yes, this means the pilot saves no time yet. That's deliberate: you're not buying speed this month, you're buying *evidence*.

Keep a simple tally: outputs right as-is / right after minor fix / wrong. Weekly, review the log with the team actually doing the checking and ask two questions: *What did it get wrong, and would our approval rules have caught it if nobody was looking?*

The mistake to avoid: quietly relaxing review mid-pilot because "it's been fine." The pilot's whole value is finding the failure shape while the net is still up.

WEEKS 11–13 — JUDGE IT LIKE A PROBATION REVIEW

End of quarter, three findings on one page for leadership:

- 1. The numbers:** accuracy rate, exceptions caught, hours actually shifted.
- 2. The failure shape:** *how* it fails (weird formats? ambiguous cases?) and whether the approval gates caught every consequential miss. (They should have. If they didn't, fix the gates before anything else.)
- 3. The verdict:** expand the role, keep it as-is, or let it go. All three are wins — a cheap, contained "no" beats a company-wide "maybe" every time.

If it passed: promote it gradually — reduce review on the proven categories, keep approvals on everything consequential, and start the next agent's job description. You now have something almost no competitor has: **a working, documented management system for AI** — the thing every future hire plugs into.

Slingr builds custom business software where AI agents work exactly this way — scoped permissions, wired-in approvals, full logs, real off switches — inside systems you own. We'll happily pressure-test your week-2 job description on a call.
slingr.io