

The COO Guide to AI Value

Your growth plan assumes people you cannot hire. Here is the operations version of the answer: your jobs, your ceiling, the plays that grow output without growing headcount, and the rhythm that keeps it honest.

2026 Edition · A role guide from *The 2026 AI Strategy*, the ebook

Delivered with  **xsparks.ai**
ACCOUNTABLE FROM IDEA TO OUTCOME

*This guide assumes the flagship ebook, *The 2026 AI Strategy* (ask; it is free). The ebook carries the full argument. This cut carries what a COO owns, and repeats only what it must.*

WHY THE COO GETS THIS CUT

Because the two numbers that define 2026 operations do not reconcile, and you are the one standing between them. Roughly three in four employers cannot fill the roles they need, near a decade high (ManpowerGroup, 2025), and a single professional hire takes about 83 days and about \$5,475 to land, by industry benchmarks (SHRM, 2025). Meanwhile McKinsey estimates 60 to 70% of the hours in a knowledge-work day go to activities AI can now perform (McKinsey, 2023). Your demand plan assumes capacity the labor market will not sell you, while most of the capacity you need sits inside your own walls, spent on routine work.

You are not short on people. You are short on the routine work leaving their desks.

The operating read compresses to one sentence: **your operation runs a model where every unit of output costs a unit of human time, and the model, not the team, is the ceiling.** The best-controlled study to date found workers with an AI assistant completed 14% more work per hour, and the lift reached

34% for less-experienced staff (Brynjolfsson, Li & Raymond, QJE, 2025). The capacity exists. It is trapped.

The precedent. Before the electronic spreadsheet, re-running a forecast meant a room of clerks and days of arithmetic, so companies asked fewer questions. When the spreadsheet arrived, one analyst did what a floor did, and firms did not shrink their finance functions; they asked a hundred times more questions, and the clerks who learned the tool became the analysts. Capacity per person is what changed. Your operation is holding the same offer.

YOUR JOBS, STATED PROPERLY

The outcome lists are candidates from field observation; ranking them against your own operation is the point.

THE JOB	THE OUTCOMES THAT MEASURE IT	WHERE IT CONNECTS TO THE FLAGSHIP
Complete the period's work during demand peaks	Maximize output per person; minimize cycle time per unit of work; minimize overtime hours per peak	The Capacity Ceiling
Fulfill each order as promised	Minimize time from sale to the promised result; maximize the share of work right the first time	The Software Tax + Deliver stage
Keep the operation visible while it runs	Minimize time to detect a problem; maximize exceptions escalated before the customer notices	Flying Blind
Retain the operation's know-how when a person leaves	Minimize time for a new hire to reach solid output; minimize the knowledge lost at a departure	The Capacity Ceiling + Moat Inversion

How your peers rank these. The surveys give you a prior. The labor constraint is structural: manufacturing alone may need up to 3.8 million new workers by 2033, with 1.9 million skilled openings at risk of going unfilled, and 2.8 million of those openings driven by retirements (Deloitte / Manufacturing Institute, 2024); 65% of manufacturers named talent their top challenge (NAM, 2024). That retirement number is why the fourth job on the list deserves more attention than it usually gets: 72% of frontline managers are not confident their company can retain the knowledge that leaves when experienced workers retire (Workplace Intelligence). The market ranking is the prior; your ranking against your own floor is still the exercise.

YOUR PRESSURE POINTS, IN ORDER

First: the Capacity Ceiling, because it caps the plan

The flagship carries the full chapter. The operations cut: the ceiling is not effort, and it is not the team. It is the share of every role's day consumed by work that should already be a system: data entry, routing, status chasing, reconciliation, re-keying between systems. Every one of those hours prices a unit of output at a unit of human time, which is why the growth plan keeps resolving to a hiring plan, and why the hiring plan keeps failing against an 83-day, decade-tight labor market.

Second: the Software Tax, because it is where the hours actually go

Workers toggle between applications roughly 1,200 times a day and lose about four hours a week just reorienting (HBR, 2022); around 60% of the day goes to work about work (Asana). For a COO this is not a finance abstraction. It is the reason your best coordinator spends her afternoon moving the same order through four systems, and it is the first place the trapped capacity comes back.

YOUR PLAYS FROM THE MENU

PLAY	WHAT IT IS	WHAT IT RETURNS
1. Agents on the volume work	In bounded, instrumented workflows, agents take the data entry, routing, drafting, and reconciliation; your people handle the exceptions	The fastest capacity recovery available; the twenty-step chain becomes a two-approval chain
2. A copilot per role	Every coordinator, dispatcher, and analyst gets senior-grade support in the flow of their work	Cycle time down, consistency up, and the biggest lift lands on your least-experienced people
3. Demand planning that pre-stages the peak	A forecasting agent that sees the surge coming and stages the work before it arrives	Peaks handled without overtime or temp hiring
4. The apprenticeship copilot	Your senior team's judgment, captured and put in front of every new hire	Ramp time falls; the retirement wave stops being a knowledge cliff

A sequencing note in your language: play 1 is where the return concentrates, and it should be one workflow, end to end, before it is ten workflows shallowly. Play 4 is the one operations leaders defer and should not: by industry estimates the knowledge walking toward retirement in your senior ranks is the single least-recoverable asset in the operation, and capturing it is a by-product of the rebuild if you sequence it in, and a crisis project if you do not.

THE WORKFLOWS AND THE AGENTS

The plays above are the strategy. Here is where they land on the floor: the specific workflows, and the agent that runs each. Rows marked *integration-heavy* need clean underlying data first; everything else is proven and reachable at mid-market scale as of this writing.

WORKFLOW	THE AGENT OR COPILOT	WHAT YOUR TEAM DOES INSTEAD	THE OUTCOME IT MOVES
Order intake and entry	An intake agent turning emails, PDFs, and photos into validated, entered orders	Coordinators handle exceptions and relationships, not re-keying	Cycle time; entry error rate
Freight and operations documents	A document agent reading bills of lading, proof-of-delivery, and RFQs, extracting fields, matching to shipments and invoices	The back office works disputes, not data entry; straight-through rates run high where deployed	Labor-hours per shipment; billing disputes
Scheduling and dispatch	A dispatch copilot proposing the day's schedule; a person approves and adjusts	Dispatchers manage the exceptions, not the grid	Jobs per crew; time to result
Quality inspection	A vision agent inspecting 100% of parts at line speed, flagging defects and learning new ones (three in four manufacturers are deploying or planning this, by industry surveys)	Quality engineers investigate causes; the line stops sampling	First-pass yield; escapes to customers
Equipment maintenance	A predictive-maintenance agent reading sensor and history data, scheduling intervention before failure (the most-validated industrial AI category: adopters report large unplanned-downtime reductions)	Maintenance plans work instead of fighting fires	Unplanned downtime; maintenance cost

WORKFLOW	THE AGENT OR COPILOT	WHAT YOUR TEAM DOES INSTEAD	THE OUTCOME IT MOVES
Inventory and replenishment	A demand-forecast agent driving reorder points, safety stock, and allocation	Planners set policy and handle the exceptions	Stockouts; working capital in inventory
Production scheduling	A scheduling agent sequencing jobs against machine, labor, and material constraints, re-planning on disruption (integration-heavy: needs clean routings and BOMs first)	Schedulers manage trade-offs, not spreadsheets	Throughput; on-time delivery
Invoice-to-job reconciliation	A matching agent tying invoices, jobs, and parts	The team works true discrepancies only	Labor-hours per transaction
Status chasing and customer updates	A status agent keeping customers and the floor informed automatically	Coordinators stop spending afternoons on "where is my order"	Time to detect; inbound status calls
Floor questions and troubleshooting	A frontline knowledge agent trained on your equipment manuals, SOPs, and job history, answering technicians on the spot. One of the best low-integration plays available	Technicians get the answer at the machine, not back at the desk; seniors stop being the human search engine	Time to resolve; senior interruptions
New-hire ramp	The apprenticeship copilot carrying your senior team's captured judgment	Seniors coach the hard cases, not the basics	Ramp time to solid output

Every row runs the series pattern: bounded, instrumented workflows, agents on the defined operating work, humans governing the seams. On the floor the seams are anything that ships, commits a promise date, or touches a customer: those carry a human approval until the evidence says otherwise, and only then loosen.

YOUR SCOREBOARD

The flagship's scoreboard mechanic applies with these rows: leading monthly in the operating meeting, lagging quarterly, two consecutive wrong-way months names the stall.

LEADING INDICATOR (MONTHLY)	TODAY'S VALUE	THE LAGGING METRIC IT PROTECTS (QUARTERLY)
Output per person		Margin; growth delivered without hiring
Cycle time per unit of work		Revenue kept (the first renewal rides on delivery speed)
Share of work right the first time		Cost to serve; customer satisfaction trend
Time to detect an operational problem		The cost of the problems you did not catch
Unplanned downtime hours		Throughput; maintenance cost; delivery promises kept
Ramp time to solid output		Capacity resilience; the cost of turnover

THE REBUILD DISCIPLINE, ON THE FLOOR

The flagship's realistic path applies in full; here is the operations half of it, the part that lives or dies on your floor.

- 1. One workflow, end to end, beats ten copilots.** Pick a flagship workflow: visible enough to matter, contained enough to finish, owned by a supervisor who wants it to work. The first workflow decides the floor's belief in everything after it.
- 2. Map today before building tomorrow.** Walk the workflow as it actually runs, not as the SOP says it runs. The gap between the two is usually where the return is.
- 3. People on the seams, from day one.** Every agent output that touches a customer or a dollar gets a human review gate at the start. Loosen the gates on evidence, never on enthusiasm.
- 4. Baseline, then measure weekly at first.** Cycle time, error rate, hours returned. The floor trusts numbers it can see move.
- 5. Say what happens to the people, early and specifically.** The first work the agents take is the work nobody wanted. Your team moves up a level: exceptions, customers, judgment calls, and the new work the rebuild creates. Lead that story before the rumor mill writes a worse one. The fuller workforce treatment is in the flagship, and the CHRO cut of this series carries the execution playbook.

Grow the output. Not the headcount.

Most of the capacity you need is already on your payroll, trapped in routine work. AI gives it back.

WHERE TO START

The capacity read, built from your numbers.

From a few inputs (the roles whose day is spent on routine work, rough headcount, and how much of the day it consumes) we build a directional estimate of the capacity trapped in your operation and what an AI-native rebuild frees. Built from your numbers, yours to keep either way. Reply and we will send it. For the full argument, ask for **The 2026 AI Strategy**, the full ebook.



Cosmo Mariano

Your AI Value Coach · Chief Client Outcomes Officer, XSparks · cosmo@xsparks.ai



Cosmo Mariano, Your AI Value Coach · The COO Guide to AI Value · 2026 Edition

A role guide from The 2026 AI Strategy, the ebook. Frameworks: Mariano IP · Delivered by XSparks, Global AI Transformation Firm · No AI Theater. Accountable from Idea to Outcome.

Sources cited inline: ManpowerGroup (2025); SHRM benchmarking (2025); McKinsey (2023); Brynjolfsson, Li & Raymond, QJE (2025); Harvard Business Review (2022); Asana, Anatomy of Work; Deloitte / Manufacturing Institute (2024); NAM Manufacturers' Outlook (2024); Workplace Intelligence. Composite figures are illustrative and directional. © 2026 Cosmo Mariano.