PITCREW

01 DECIDE 02 GROUND 03 CALIBRATE 04 PROVE 05 OPERATE

LOWER OPERATING COST DEFENSIBLE OUTPUT CAPACITY WITHOUT HIRING

Five Steps to Governed AI

A discipline for firms where the cost of being wrong is higher than the cost of being slow.

GOPITCREW.COM

Five Steps to Governed AI

A discipline for firms where the cost of being wrong is higher than the cost of being slow.

EDITION — First edition, 2026

CONTACT — sales@gopitcrew.com · gopitcrew.com

Buying AI is easy. Collecting its return is a discipline.

This book is that discipline: five steps that turn AI spending into measurable returns. The pages that follow show the steps, the operations they fit, and what each one is worth to the firm running it.

What ungoverned AI costs. What governed AI returns.

THE COST

$50K–$150K	A year in fee errors, per firm, surfacing downstream when the money is already out the door. Every hand-run operation carries its own version of this number.
40 hours	Of senior time each time someone asks “how do you know?” and the answer has to be reconstructed from email and memory.
6 months	To rebuild judgment when a key person leaves, because the rules were never written anywhere but their head.

THE RISK

95%	Of AI pilots never return their investment. Not a capability ceiling: a missing proof step, holding the money at the pilot stage.
14%	Of CEOs have defined the P&L impact of their AI initiatives, while more than half name that missing link as their key barrier. Unproven output is unaccountable spend.
Liability	Every output no one can account for is carried risk: for the answer, and for every hour spent establishing whether an error happened.

THE VALUE

Lower cost	Errors caught before they reach a client. Nothing to refund, nothing to claw back.
Defensible output	Proof produced as a byproduct of the work, shared in the time it takes to be asked for it.
Capacity	Growth without proportional hiring. The discipline scales; headcount does not have to.

The five steps in this book are the move from cost and risk to value.

FIGURES: SEE METHOD AND SOURCES, PAGE 9

Two ways the same AI spend behaves.

The numbers on the previous page split firms into two groups. Which group you are in depends on how the spend behaves.

UNGOVERNED AI

An expense

- Tools at the edges, demos in the middle
- Output nobody trusts with real operations
- The cost base it was bought to change, unchanged
- Renewed every year. Returns more demos

GOVERNED AI

An investment

- The same models, wrapped in your firm's own rules
- Tested against your own history before it runs
- Work you can bill, defend and build on
- Returns collected every quarter, on the operating line

The return shows up in four numbers: cost per unit of work, errors reaching clients, time to answer an audit, and volume handled per head. Chapter two prices each one.

Most firms that buy AI get exactly what they paid for: a system that works. What they do not get is a system they can stand behind.

That gap is where the money goes. Firms lose it in refunds, in write-offs, in senior hours spent redoing finished work, and in growth that stalls because every new client means new hires. The AI did not fail. The investment just never reached the operations where the returns are.

When the gap closes, the same spend starts paying. Work goes out right the first time. Questions that took weeks to answer take minutes. The team you have handles the growth you win. That is what AI was supposed to buy, and it is available to any firm willing to run five steps in order. This book is those five steps. It is not a product pitch. The discipline works regardless of who builds it.

But I will be direct. We built PitCrew because we believed this discipline needed a machine behind it, not just a book about it. If you want to see what it looks like running on real accounts, the last chapter tells you how.

Rishi Kulkarni

CO-FOUNDER & CEO, PITCREW

Four signs this book was written for your firm.

An error reaches a client as money.

A wrong fee, a wrong payment, a wrong claim. When mistakes are denominated in dollars, the output has to be provable, not plausible.

Someone can ask “how do you know?”

A board, a regulator, an auditor, an acquirer, a client. If that question has to survive scrutiny, a confidence score will not carry it.

The rules live in people’s heads.

Approvals, exceptions and definitions of correct that have never been written down anywhere a machine could read them.

Growth currently means hiring.

Every new account or client adds headcount, because every process is checked and approved by hand.

A NOTE ON INDUSTRIES

Wealth management or logistics, insurance or healthcare, banking or the enterprise back office: the operations differ, the failure mode does not. Wherever an output must be defensible, a fee, a payment, a claim, a contract, a compliance filing, the five steps apply unchanged.

CONTENTS

Where the numbers in this book come from.

A book about provable output should say where its own numbers came from. Each one below is research, estimate, or illustration.

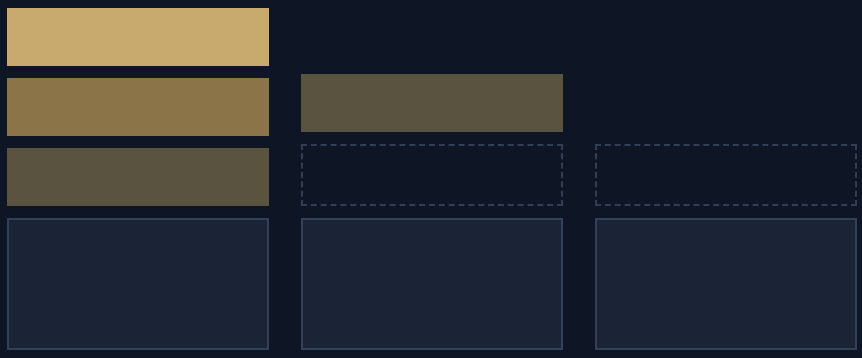
95% of AI investments never reach production. Published research. Quoted as reported.	MIT NANDA The GenAI Divide, 2025
AI costs have fallen close to 300x in two years.	Public model pricing, 2024 to 2026
The time figures on the ten operations pages. Marked INDUSTRY ESTIMATE wherever they appear. Ranges, not averages.	Practitioner interviews and firm process reviews
Fee errors cost advisory firms $50K to $150K a year. A range reported across operations leaders at mid-size firms. Treat it as an order of magnitude, not a benchmark.	Industry estimate
Accounts, names, ledger balances and rule numbers. Composites. Every example in this book is constructed. No client, firm or transaction is real.	Illustration
Half of CEOs cite linking AI to the P&L as a key barrier; 14% have defined that impact. Published survey research. Quoted as reported.	BCG AI Transformation CEO Survey, 2026

WHAT THIS BOOK IS NOT

A benchmark study. We did not measure firms against each other, and we make no claim about how many have adopted any of this.

CORRECTIONS

If a figure here is wrong, we would rather know. Write to sales@gopitcrew.com and the next edition will carry the correction.



Same AI, Different Outcome

The cost of AI fell 300x in two years. The cost of getting it wrong stayed exactly the same. The gap between the two is where firms gain an advantage or lose money.

The cost of intelligence fell 300x. The cost of being wrong did not.

AI is now available to every firm at near-zero cost. AI costs have fallen close to 300x in two years. When the technology is identical, it stops being an advantage.

300×

fall in AI cost in two years. The same intelligence, at the same price, to every competitor.

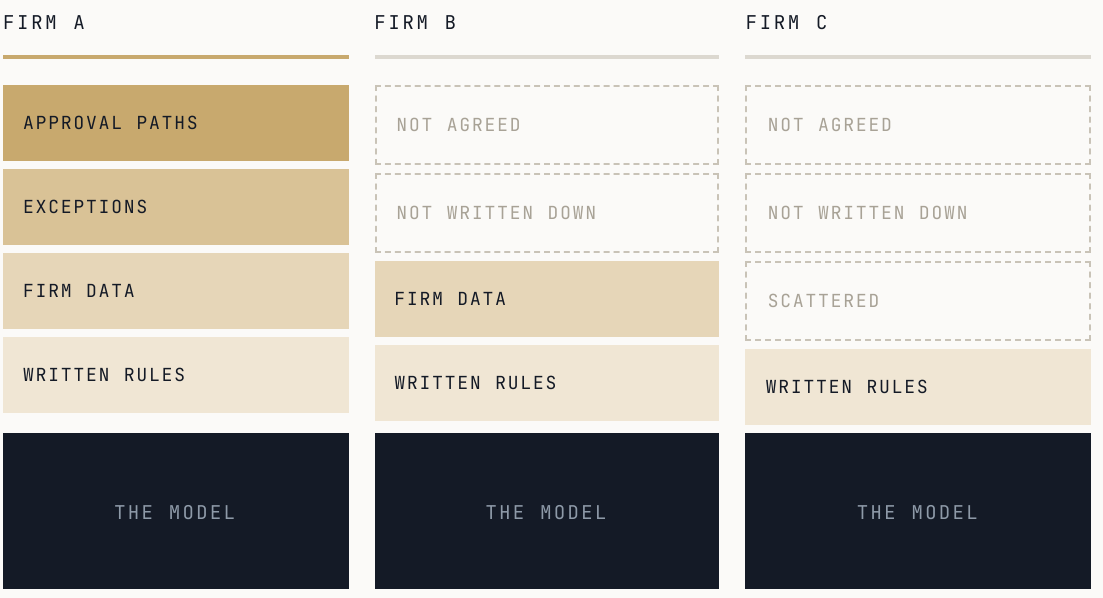
The question every leader is asking: what do we have that the firm next door does not?

The answer is your firm's judgement: The values, the policies, the goals, the relationships. It lives in your people and your culture, not in any model's training data.

Two problems follow. Most of it was never written down in a form a machine can use. And even when the machine uses it correctly, you have no way to prove that it did. Unwritten judgment and unproven output: that combination is what keeps AI investments from returning anything. This book sets out the discipline that fixes both.

FIGURE 1.1

Three firms, one identical model. The gap is the operating line.



The base layer is rented and identical. Everything above it, the rules, the exceptions, the sign-offs, is the only part that produces a different outcome. Firm C has a model and nothing to put it into. Same investment, no return. And three people know how it actually works: when they leave, the firm starts from zero, the next hire takes six months to reach the same judgment, and errors compound in the gap.

AVAILABLE TO EVERY COMPETITOR

Reasoning, language, code, speed, scale

YOURS ALONE, IF WRITTEN DOWN

Your rules, your exceptions, your sign-offs, your proof

The exception that costs $200K

Written in a 2019 policy memo. Operations and reporting read it differently. Both are defensible. The AI picks one, and the wrong client gets the wrong fee.

The approval that delays a client two weeks

Depends on the amount, the account type, and who is available. Nobody wrote the decision tree, so new hires guess and clients wait.

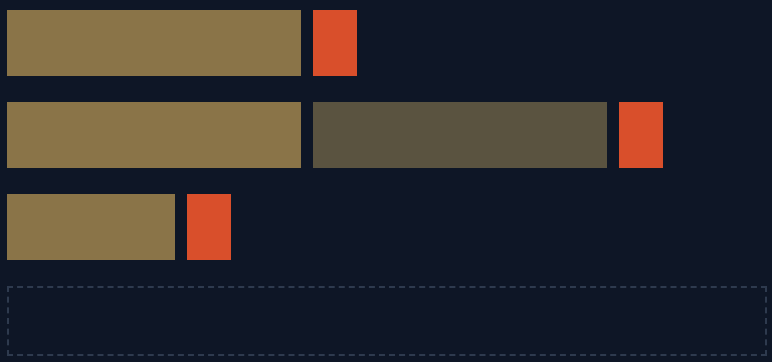
Which system wins when two disagree

No source of record designated. Each team defaults to its own, and the client gets two different numbers on two different statements.

The definition of correct that lives in four heads

They all agree when asked. None has written it down. When the first one leaves, the definition changes and nobody notices.

A better model does not fix this. These are decisions the firm has never made explicitly. Until someone makes them, writes them down and gets them signed, the firm pays for the ambiguity in errors, delays and rework every quarter.



Where the ROI Disappeared

Three ways firms invest in AI without changing the cost of operations. Every one of them stops one step short of return.

Most AI investments stop delivering return after the first demo.

Rarely because the technology is wrong. Usually because nothing governs how an output gets proven, which means nothing reaches production. Here is what each pattern looks like from inside the firm.

PATTERN 01

Spent on tools, not operations

The CFO approved six AI tools this year. Operating cost per account has not moved. Each tool does one narrow thing well, and the operations that actually cost money are untouched: billing, onboarding, reconciliation.

PATTERN 02

Spent on building, not proving

The CTO’s team has been building for eighteen months. The board still asks how anyone knows the output is right. Years of development and tens of millions invested, with no agreed method for answering that question.

PATTERN 03

Spent on demos, not production

The COO saw it work in March. It is August and nothing runs on live accounts. Someone asked how they would know an output was right, nobody could answer, and no production date was ever set.

WHERE IT SITS

AI handles tasks nobody was measured on. The operations that drive cost are unchanged.

THE RESULT

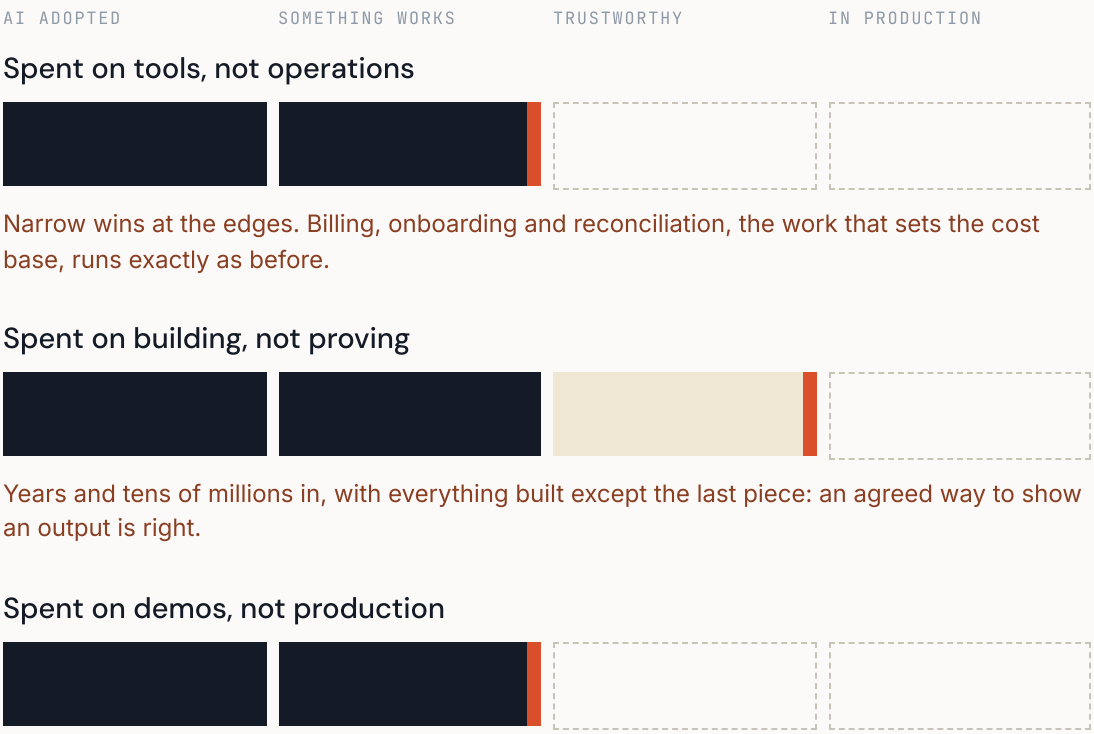
Budget spent, demos presented, and no change to operating cost or capacity.

THE GAP

Nothing between a system that works and one that returns anything. The missing step is proof.

FIGURE 2.1

Three routes, and where each one stops.



THE GAP, IN ALL THREE

Between a system that works and one that delivers return there is one missing piece: the discipline for proving the output is right. Without it every investment stops at the demo.

WHAT IS MISSING

A standard for proving the output, so the investment can reach production

THE COST

Every quarter run manually is a quarter the investment returns nothing

The AI was capable. It did not know the business, and nothing was built to prove it had followed the firm's rules. Not a technology failure. A proof failure.

WHAT WAS BUILT

Something that worked in a demo, on data someone chose.

WHAT WAS NEEDED

Proof it would be right on the accounts nobody chose: the exceptions, the edge cases, the ones that cost money when wrong.

What it costs to run without the discipline.

Five places the gap between “it works” and “we can prove it works” shows up on the operating line. Every figure here is carried every quarter, not once.

Fee accuracy	$50K–$150K	In errors a year, found at the annual review after the client has already paid.
Answering “how do you know?”	40 hrs	Of senior time per request, reconstructing six months out of email and memory.
A key person leaves	6 months	For a new hire to reach the same judgment. The process rebuilds from zero and errors compound in the gap.
The board asks for proof	Days	Senior people pulled off revenue work to assemble records nobody kept as they went.
Capacity	1:1	Every new account adds headcount. Growth costs in proportion to the work won.

CARRIED TOGETHER

A firm running these five operations by hand carries all five costs at once, every quarter, and none of them appears as a line item anyone owns.

What replaces it when the output is provable.

Same five operations, same order. The discipline does not make the work faster. It makes the answer hold, which is what removes the cost.

Fee accuracy	Zero	Errors caught before invoicing. The client never sees them, so there is nothing to refund.
Answering “how do you know?”	Seconds	An export. Produced in the time it takes to ask for it, by whoever was asked.
A key person leaves	No rebuild	The process survives the person. Written, signed and still running.
The board asks for proof	Immediate	Share the record. It assembled itself while the work was happening.
Capacity	Flat	Grow without proportional hiring. The discipline scales.

HOW TO MEASURE THE RETURN

Most firms measure AI by activity: hours saved, tools adopted, demos shipped. Activity is not return. The return is four numbers: cost per unit of work, errors reaching clients, time to answer an audit, volume per head. If AI is paying, all four move. Worksheet 04, at the back of this book, turns them into your firm’s numbers.

THE NUMBER EVERY BOARD HAS SEEN

95%

of AI pilots never reach production.
The technology was never
the reason.

MIT NANDA, THE GENAI DIVIDE, 2025

01

02

03

04

05

From Pilot to Operating Leverage

Five steps. Each one removes a specific cost or risk. Skip one and the investment stays where it is: a demo that never reaches the work that costs money.

A discipline, not a product.

That sequence has a name only because the industry needs a standard for it. Five steps, each closing a failure mode that costs real money, in an order that cannot be rearranged because each step depends on the one before it.

It works with any model, any technology stack, any team, and it applies whether you build the AI yourself or someone builds it with you. We are sharing it because a standard kept private is not a standard. If your firm proves its output and the firm next door does not, you are not in the same competitive position, whatever model you both run.

ANY MODEL, ANY TEAM

The discipline is the standard. The technology underneath is a choice.

CUMULATIVE

Each step depends on the one before it. Skip one and the return from every step after it drops.

FIGURE 3.1 / THE FIVE STEPS

How an AI investment becomes operating leverage.

IN · AI INVESTMENT — Cost centre. Unproven.

01 Decide

Agree what the AI should do, in plain language, signed by whoever owns the process.

SKIP IT: REWORK COSTS MORE THAN THE BUILD

02 Ground

Connect it to your firm's live data, documents and policies, with a named source of truth.

SKIP IT: ERRORS REACH CLIENTS AS INVOICES

03 Calibrate

Run it against your hardest real scenarios and compare it to what your team actually did.

SKIP IT: THE FIRST FAILURE IS IN FRONT OF A CLIENT

04 Prove

Check every output against your rules. A separate system. Pass or fail, with the rule named.

SKIP IT: THE INVESTMENT CARRIES LIABILITY, NOT RETURN

05 Operate

Your team approves what executes. The record assembles itself as a byproduct.

SKIP IT: NO RETURN ON THE TIME SAVED

OUT · OPERATING LEVERAGE — Proven. Creating Value.

FIGURE 3.2 / VALUE CREATED, RISK REMOVED

What each step is worth, and what it takes off the table.

		VALUE TO THE FIRM	RISK ELIMINATED
01	Decide	A written agreement that survives a personnel change. No rework.	No one can say afterwards what the process was meant to do.
02	Ground	Answers drawn from your own records, at the moment they are asked for.	Confident output invented from material that was never yours.
03	Calibrate	A measured rate of agreement with the people who do the work today.	The first hard case arrives in front of a client instead of in a test.
04	Prove	A formal check that your rules held on this specific output.	Liability for output no one can account for. The cost of being wrong.
05	Operate	Proof produced while the work happens. Capacity grows without hiring.	Senior people rebuilding six months of history out of email.

WHY THE ORDER HOLDS

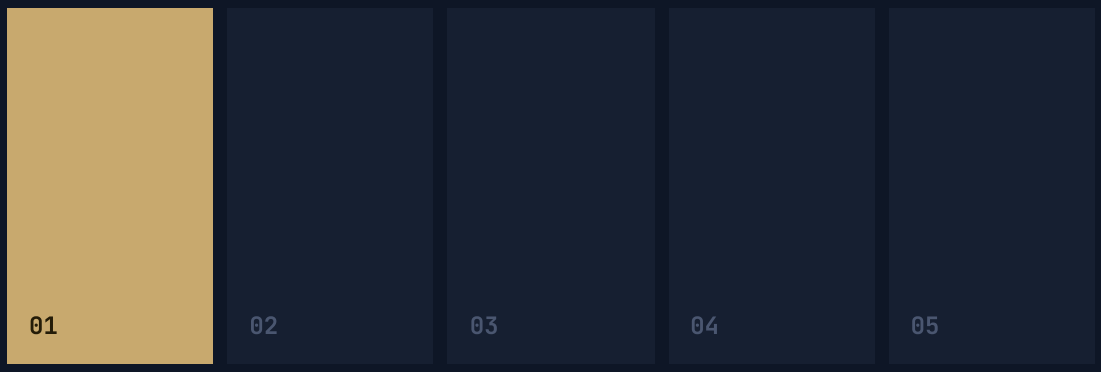
The steps are cumulative. Proving an output against rules nobody agreed proves nothing. Skip one and every step after it returns less.

HOW TO READ THE NEXT FIVE CHAPTERS

Every step is set out the same way.

Each chapter opens with what it costs to skip the step, then shows what the step is and what good looks like.

What the step is	What it actually is, in one paragraph, with no technology in it.
What it costs to skip	The dollar cost, the time cost, or the risk you carry without this step.
What good looks like	The document or behaviour you should be able to point at when the step is done.
At your firm	A situation you have lived through. The cost is named.
Skip this step	What breaks, and where it shows up.



Decide

Agree what the AI should do before anything is built. In plain language. Signed by whoever owns the process.

Skip it and the AI builds to one person's assumption. Fixing it costs more than doing it right for the first time.

WHY IT MATTERS

Without a definition, the agent does what someone assumed the process was.

Two people each assume the other owns a step. An exception nobody wrote down never reaches the build. The agent then does exactly what it was told, which is not what the process is.

WHAT GOOD LOOKS LIKE

One page stating what triggers the process, which rules apply, which exceptions exist, and what a correct output looks like, specific enough that two people reading it would build the same thing. It is signed by whoever owns the process: operations, finance, compliance. The signature turns a set of assumptions into a decision someone has made.

It is not a requirements document, and nothing in it should name a system or a model.

IT IS	IT IS NOT
One page. Plain language. Signed. Written before anything is built.	A fifty-page PRD. A backlog. A diagram nobody signed.

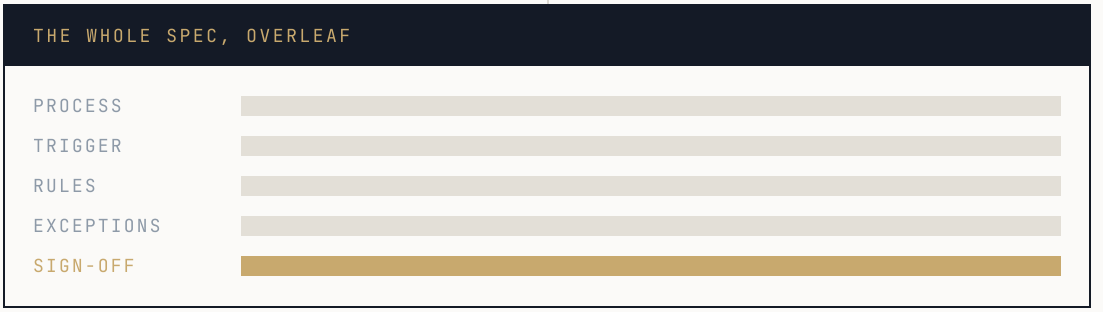


FIGURE 4.1

What the agreement looks like

PROCESS SPEC v1.0 · SIGNED

PROCESS

Vendor invoice approval and payment

TRIGGER	Invoice received and matched to a purchase order.
RULES	Invoices under $5K auto-approve if the PO matches exactly. $5K to $10K need department manager sign-off. Above $10K needs VP Finance. Variance above 2% from the PO routes to the controller.
EXCEPTIONS	Recurring subscriptions with annual POs skip per-invoice matching. Emergency purchases approved by the COO bypass threshold routing.
CORRECT	Paid to the right vendor, at the right amount, with the right approval, within payment terms.

SIGN-OFF

M. Torres, VP Finance

WITHOUT IT

Both teams believed the other owned the threshold rule. Neither was wrong, because nobody had written it down.

IN YOUR FIRM

Your AI approved 340 vendor invoices last month. Nobody told it what “approved” means.

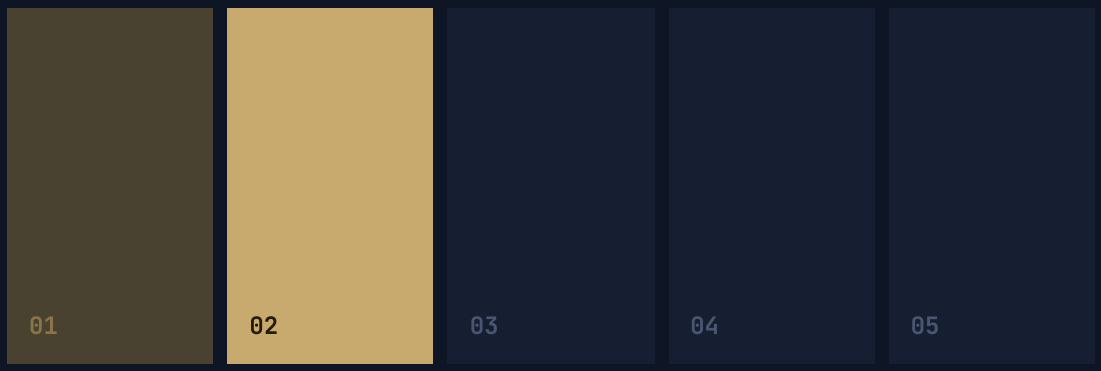
Finance set a $10K threshold for manager sign-off. Procurement set $5K. The AI was built on Finance’s rule, so for two months invoices between $5K and $10K went through unapproved.

WHO ASSUMED	Engineering assumed Finance owned the approval logic.
WHAT WAS MISSED	Procurement’s $5K threshold was never written into the spec.
WHO NOTICED	An external auditor, during the quarterly review.
WHAT IT COST	$287K needing retroactive sign-off, and an escalation to the board.

The AI did exactly what it was told. Nobody had agreed what to tell it.

SKIP THIS STEP

The agent builds the wrong thing. Nobody agreed what “correct” means.



Ground

Connect the agent to your firm's actual data, documents and policies. This is what it references instead of guessing.

When systems disagree, someone has to decide which one wins. Skip that and the difference reaches a client as an invoice.

WHY IT MATTERS

Ungrounded, the agent answers from someone else’s records.

Ungrounded, the agent answers from what the AI agent absorbed in training: a fee schedule belonging to some other firm, a reading of a rule that is not your reading, occasionally a document that never existed. What your firm knows sits in the ERP, the AP system, the contract files and the policy drive. The agent has to read those at the moment it is asked, not from anything cached at build time.

WHAT GOOD LOOKS LIKE

A written map of which system holds what, and a designated source of record for every field where two systems could disagree. Someone senior signs that designation. Until they do, every answer downstream carries the ambiguity forward, and you will not be able to tell which answers were affected.

GROUNDED IN

Live systems, current documents, your written policies.

NOT GROUNDED IN

Training data, last quarter’s export, an assumption about market practice.

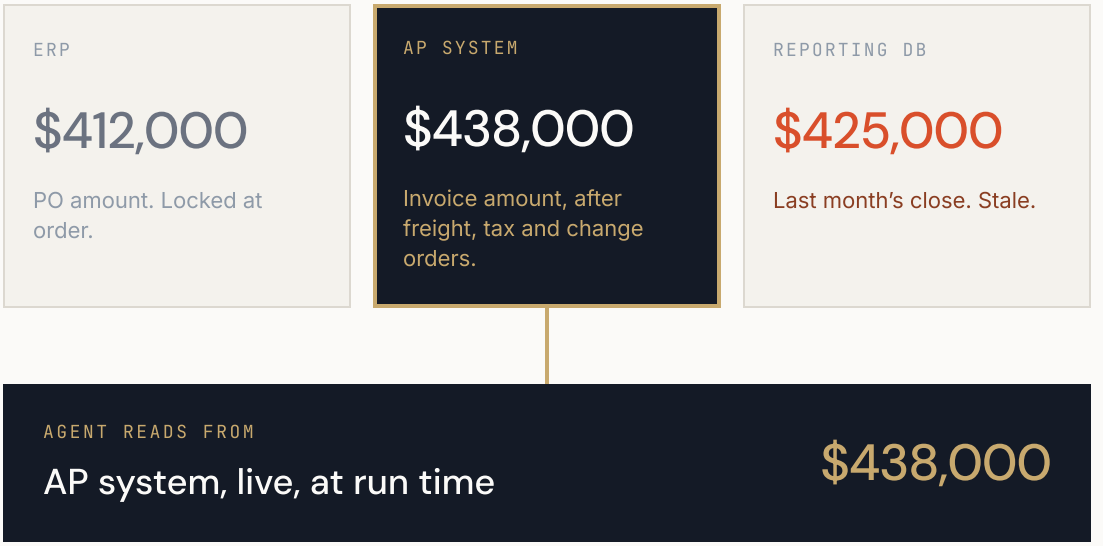
ONE OF THEM WINS

ERP	AP system	Reporting
REFERENCE	TRUTH	DOWNSTREAM

FIGURE 5.1

Three systems. One vendor. Two answers.

Vendor 40-118843, amount payable, same close date.



Until the winner is named and written down, the agent pulls from whichever system it queries first, and the difference reaches the vendor as an overpayment.

THE DECISION, WRITTEN DOWN

AP system	Invoice amount, adjusted	TRUTH
ERP	Purchase order amount	REFERENCE
Reporting DB	Monthly close snapshot	DOWNSTREAM

IN YOUR FIRM

Your ERP says you owe the vendor $412,000. Your AP system says $438,000. Which number did the agent use to cut the check?

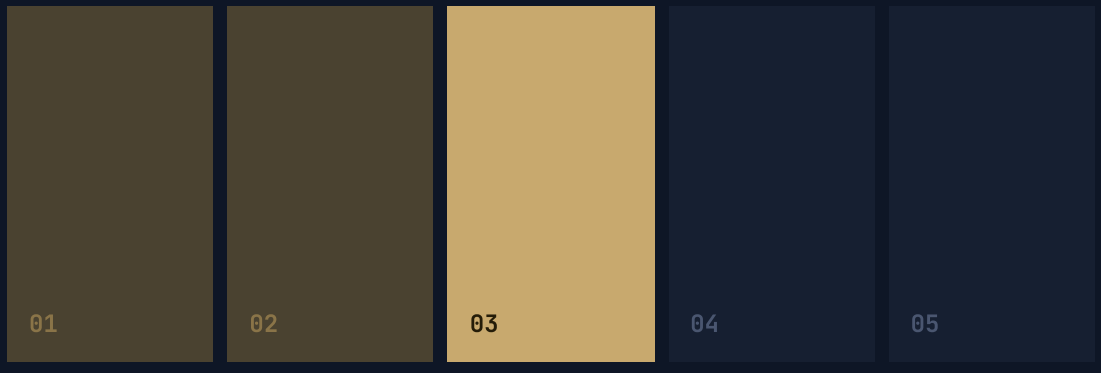
Both are correct. The ERP books the purchase order amount; the AP system books the invoice after freight, tax and change orders. The delta is $26,000 sitting between two definitions of what you owe.

THE SPREAD	$26,000 per vendor, across 14 vendors. $364,000 in unreconciled payables.
THE DECISION	Never made. No system was named the source of truth for payment.
THE COST	Two overpayments totalling $51,000 clawed back over 90 days. One vendor relationship damaged.

The number was defensible in one system and wrong in the check. Nobody could say afterward which system had been used.

SKIP THIS STEP

The agent hallucinates. It uses AI memory instead of your actual records.



Calibrate

Run it against real scenarios from your firm and compare the results to what your team would have done.

If it cannot match your team on the hardest five, it is not ready for the other five hundred. The alternative is that the first error reaches a client.

WHY IT MATTERS

Sample data is not your book of business.

Sample data is clean because someone cleaned it. Your book is not. Untested, the first real exception surfaces in front of a client rather than somewhere you can still change the answer.

WHAT GOOD LOOKS LIKE

Pull real cases from the past year: the exceptions, the unusual account structures, the ones that took someone an afternoon. Run the agent against them. Compare each answer to what your team actually did at the time. Use the hard ones: the easy cases will agree and tell you nothing. Where the two differ, one of them is wrong, and establishing which is the work.

85–95%	Agreement with your team. It is ready.
Below 85%	Tune it. Something in the spec or the grounding is wrong.
Above 95%	Investigate the disagreements. It may be right where your team was inconsistent.

The point of the exercise is not the score. It is the conversation the disagreements start.

FIGURE 6.1

Your five hardest cases, side by side.

Real expense reports from the past year. The agent's answer against what the controller actually did.

SCENARIO	AGENT	TEAM	MATCH
Standard travel reimbursement	$1,240	$1,240	
Multi-city trip, split billing	$4,880	$4,880	
Conference with client dinner	$2,310	$2,310	
Home office kit, policy changed mid-quarter	$3,400	$3,400	
Client entertainment dinner, two attendees	$9,400	FLAGGED	

AGREEMENT

80%

Below the line. Not ready.

THE ONE THAT CAUGHT IT

The agent applied the per-event limit. The controller applies the per-head limit. That rule was never in the spec, so it was never in the agent. Caught in calibration, not in a reimbursement.

WHAT HAPPENED NEXT

01 **The exception went into the spec.** The per-head cap applies to all client entertainment. Signed by the controller.

02 **The agent was rebuilt against the amended spec.** Not patched. The spec is the source, so the change starts there.

03 **The five cases were re-run.** Along with the 195 that already passed.

AGREEMENT, SECOND PASS

100% The loop is the method. A failed test is the step working, not the step failing.

The agent matched 200 expense reports to policy. Then it approved the one your controller would have caught in five seconds.

A $9,400 dinner receipt. Policy caps client entertainment at $150 a head, and the receipt listed two attendees. The agent read it as business development and passed it against the $10,000 per-event limit.

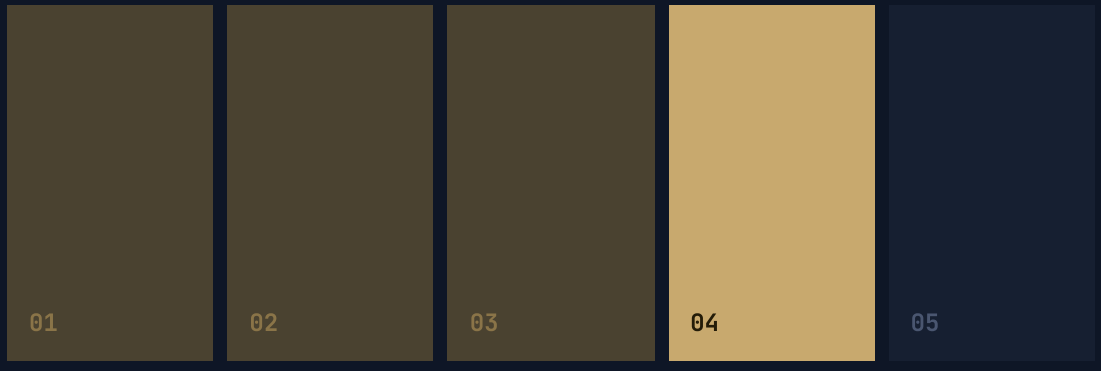
200 EXPENSE REPORTS THE AGENT MATCHED	201

WHAT THE POLICY SAYS	Client entertainment: $150 a head. Nothing above VP without CFO pre-approval.
WHAT THE AGENT MISSED	Per-head math. It checked per-event instead.
WHAT IT COST	Reimbursement paid. Flagged three months later, at the next internal review.

Report 201 was never in the calibration set. That is the only reason it was in production.

SKIP THIS STEP

Edge cases blow up live. The first failure is in front of a client.



Prove

Before any output reaches your team or your clients, it is checked against your firm's rules.

Pass or fail. With the rule named.

WHY IT MATTERS

A confidence score is not proof.

Unverified, the firm carries liability for every output. When something goes wrong the cost is not the error itself. It is every hour spent establishing whether the error happened, plus every client who now checks your work before paying. A confidence score is a property of the model, not of the answer. What survives the question is the rule, the input, the output and the link between them.

WHAT GOOD LOOKS LIKE

A separate verification layer, not the same AI grading its own homework, working from your written-down policies and returning a definitive pass or fail. On a failure you see the rule that broke and the data that caused it.

> In PitCrew Factory, this layer is powered by **AWS Automated Reasoning**, a formal verification system that checks the rules as written. It is not a second model checking the first.

PROOF	NOT PROOF
R 3.4 was applied to this input and produced this output. Pass.	The model reported 94% confidence.

THE SAME OUTPUT, TWO ANSWERS

VERIFIED

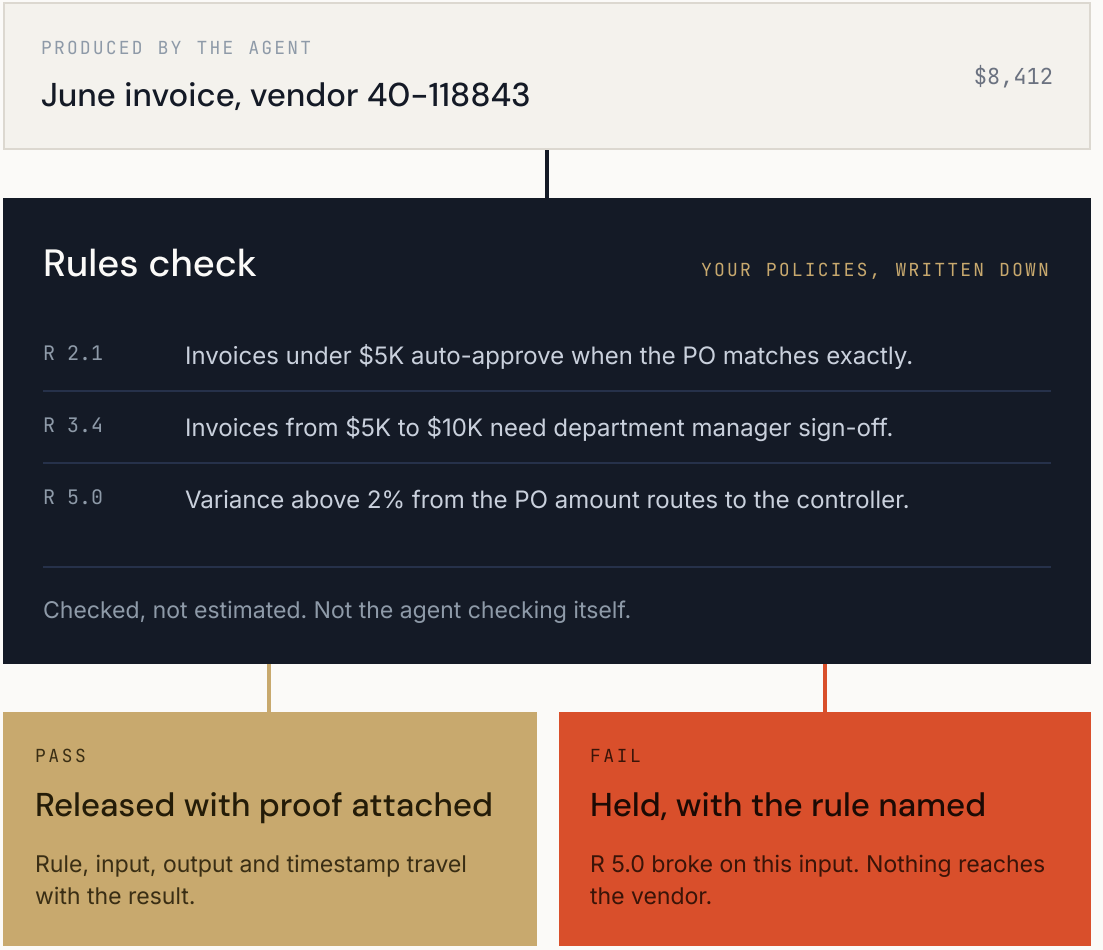
```
R 3.4 applied to this input.
Output consistent.
Pass.
```

UNVERIFIED

```
94% confidence.
Source not cited.
No rule named.
```

FIGURE 7.1

The rules check. A different system does the checking.



The gate is separate from the thing it is checking. That separation is what makes the result proof rather than an opinion.

WHAT TRAVELS WITH A PASS

THE RULE	THE INPUT	THE LINK
R 3.4, as written and approved	Invoice amount at close, timestamped	Why the two are consistent

IN YOUR FIRM

The agent reports that the wire matches standing instructions. The routing number is one digit off from the file.

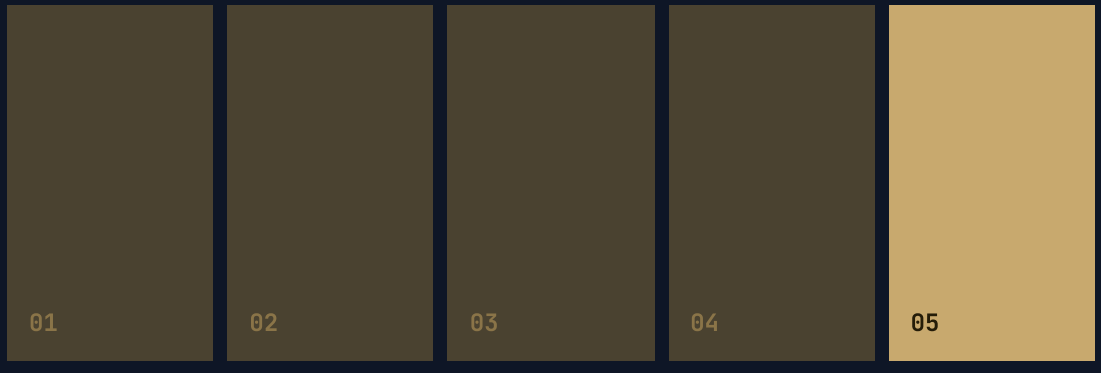
Pattern matching found a near-fit and called it a match. A rules check compares the field exactly, and catches it before $200K moves to the wrong account.

WHAT THE AI SAW	A routing number that looked right.
WHAT THE RULE SAW	A field that did not match the record, character for character.
WHAT IT SAVED	The $200,000 that would have left, plus the recall, the client call and the incident report.

The number was there. Nothing recorded which rule had produced it.

SKIP THIS STEP

You cannot prove the output is correct. A confidence score does not survive that conversation.



Operate

Your team approves what runs. The record assembles itself as a byproduct.

The proof assembles itself. Nobody types it after the fact.

WHY IT MATTERS

Otherwise you spend as much time proving the work as doing it.

An agent that does useful work and keeps no record of it leaves you reconstructing that work later from email and memory. The reconstruction costs more senior time than the work saved.

WHAT GOOD LOOKS LIKE

The agent runs the process. It surfaces exceptions for human review. Your team approves what needs approval. Every action is logged with a timestamp, the data it referenced, the rule it applied, and the decision made. When someone asks, your board, a client, a regulator, an acquirer, the proof is already assembled. You share it. No scramble.

THE AGENT DOES

The work, the logging, and the flagging of anything unusual.

YOUR TEAM DOES

The judgment calls, and the approval that puts a name on the record.

That division is what makes the trail real. The agent cannot approve its own work. Your team cannot be asked to log it by hand.

FIGURE 8.1

The trail writes itself.

Monthly vendor payment run, 847 invoices. Nobody typed any of this.

09:00:04	GROUNDED	AP system, invoice close 30 Jun, 847 invoices read
09:00:31	MATCHED	POs matched to invoices, 12 variance exceptions flagged
09:01:08	CHECKED	834 pass against R 2.1, R 3.4, R 5.0
09:01:08	HELD	13 invoices fail threshold or variance rules, routed for review
10:18:22	APPROVED	M. Torres, VP Finance, 11 exceptions cleared with note
10:18:30	EXECUTED	845 payments queued to bank
10:18:31	LOGGED	Proof sealed, exportable

WHEN SOMEONE ASKS

“Show me the 13 exceptions and who cleared them.”

ALREADY ASSEMBLED

Line four and line five. Shared in the time it took to ask.

IN YOUR FIRM

Your board asks how you know the agent did what you said it would, on every payment, for the last six months. You have the work. You cannot show the working.

With the record running, you export it: every check timestamped, every exception flagged, every approval named. None of it was written after the fact.

WITHOUT THE TRAIL	Four people pulled off current projects. Three weeks reconstructing logic from logs and memory, ending in a deck that says “we believe.”
WITH THE TRAIL	One export. Every decision, the rule it applied, the data it saw. Board satisfied before lunch.
WHO IS PULLED OFF	Nobody. The record does not live in anyone’s memory.

The work was done correctly. That is not the same as being able to show it.

SKIP THIS STEP

No record. No human approval. No proof when someone asks.

The Discipline Applied

Ten operations, laid out the same way so the pattern is visible before the detail. The vertical is incidental: swap the custodian for a carrier, the ledger for an ERP, and the five steps run unchanged.

A sample of sixty, by category.

Each Skill is a process that has already been through the five steps once.

- The ten marked are worked through in full on the pages that follow.

DATA & RECONCILIATION

- **Standing Instruction Recon**
  Platform against custodian, daily
- **Portfolio Drift Check**
  Holdings against policy bands
- **AUM Reconciliation**
  Feed against ledger at close
- **Cash Position Monitor**
  Low balance and sweep alerts
- **Data Quality Audit**
  Field completeness by system

BILLING & FEES

- **Fee Billing Validation**
  Agreement terms against invoice
- **Commission Reconciliation**
  Payout against production
- **Invoice Matcher**
  Line item to source record
- **Billing Exception Detector**
  Variance with root cause
- **Fee Tier Auditor**
  Breakpoints and household rollup

ONBOARDING & OFFBOARDING

- **Account Opening Filler**
  Signed docs to custodial form
- **Client Doc Collector**
  Requests, receipts, reminders
- **Onboarding Form Validation**
  Blocks until every rule passes
- **Offboarding Checklist**
  Closure steps and confirmations
- **Transfer Coordinator**
  In-kind position tracking

CONTROLS & COMPLIANCE

- **IPS Drift Compliance Check**
  Drift, bands, signature currency
- **KYC Collection and Verification**
  Collection and escalation
- **ADV Update Monitor**
  Change detection and filing dates
- **Wire and ACH Validation**
  Standing instructions and limits
- **Regulatory Calendar**
  Obligations by date and owner

DOCUMENT INTELLIGENCE

- **PDF Extractor**
  Structured fields from any form
- **Agreement Parser**
  Terms, rates, exceptions
- **Statement Reader**
  Positions and transactions
- **Tax Return Analyzer**
  Schedules and carryforwards
- **Policy Bound Setup**
  Record, renewal, welcome, filing

60+

SKILLS IN THE CATALOG

The catalog grows as firms contribute the processes they have already disciplined.

IPS Drift Compliance Check

Know before your client does that a portfolio has drifted past the policy that governs it.

WHO FEELS IT · CHIEF COMPLIANCE OFFICER

WHAT IT TAKES TODAY

40+ HOURS QUARTERLY

INDUSTRY ESTIMATE

A compliance analyst opens each IPS, transcribes target allocations into a spreadsheet, pulls holdings from the portfolio system, and calculates drift per asset class for every account under management.

WHAT THE AGENT DOES

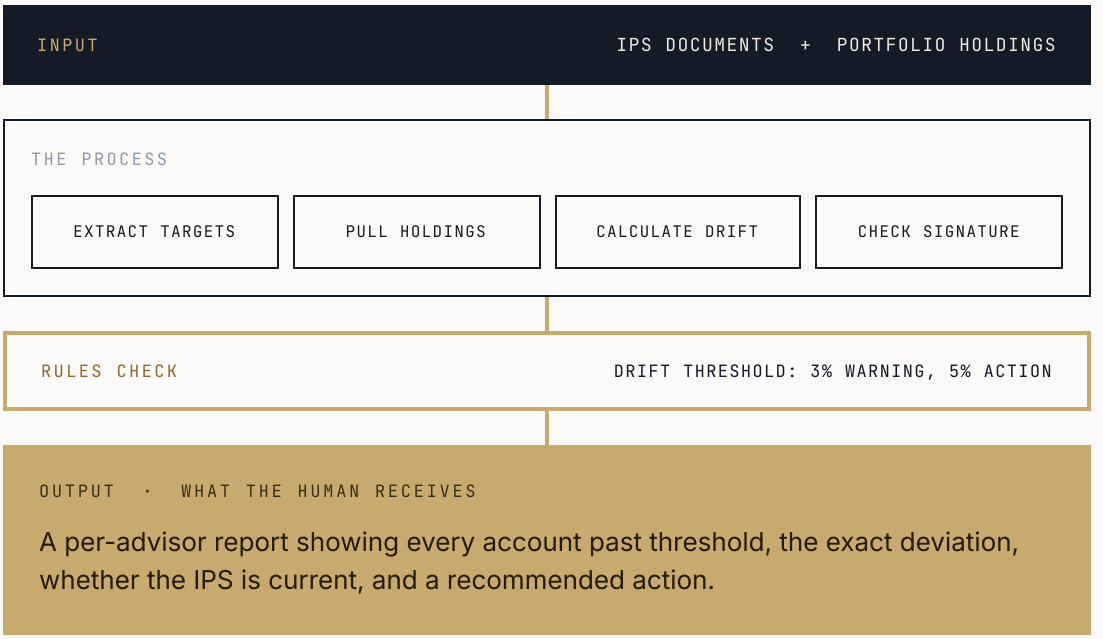
01 Reads each IPS document, pulling target allocations, tolerance bands and restrictions into a structured record.

02 Queries the portfolio management system for current holdings and calculates actual allocation per asset class.

03 Compares actual against target and flags any position past the stated drift threshold.

04 Checks for low cash and confirms the IPS carries a valid signature and a review date inside twelve months.

05 Produces a per-advisor exception report with the drift amount and the recommended rebalancing action.



FIVE STEPS MAPPED

01 DECIDE	02 GROUND	03 CALIBRATE	04 PROVE	05 OPERATE

DECIDE	Which IPS fields are rules: target, band, restriction.
GROUND	The portfolio system API is the single source for holdings.
CALIBRATE	Fifty accounts with pre-calculated drift confirm the math.
PROVE	Compliance reviews flagged exceptions before any trade.
OPERATE	Weekly, on a schedule, exception-only output.

RESULT

~~40+ hours quarterly~~ — A 15-minute review of flagged exceptions

Fee Billing Validation

Billing errors caught while they are still a variance, not after the client has read the invoice.

WHO FEELS IT · CONTROLLER / CFO

WHAT IT TAKES TODAY

3–5 DAYS PER BILLING CYCLE

A billing analyst cross-references each client's agreement fee schedule, custodian-reported AUM and the generated invoice, line by line, every quarter.

INDUSTRY ESTIMATE

Fee errors are estimated to cost advisory firms $50K to $150K a year in clawbacks and attrition.

WHAT THE AGENT DOES

01 Extracts fee schedule terms from advisory agreements: tiered rates, breakpoints, caps, household aggregation.

02 Pulls actual AUM from custodian statements or data feeds for the billing period.

03 Calculates what the fee should be, applying the agreement terms to the actual AUM.

04 Compares the calculated fee against the invoice the billing system produced.

05 Flags every discrepancy with the dollar variance, the root cause and a route to the controller.

INPUT — AGREEMENTS + CUSTODIAN AUM + INVOICE

THE PROCESS

EXTRACT TERMS | APPLY TO AUM | RECALCULATE | COMPARE

RULES CHECK — FEE SCHEDULE RULES: TIER, BREAKPOINT, HOUSEHOLD

OUTPUT · WHAT THE HUMAN RECEIVES

A billing exception report listing every account where the invoice does not match the calculated fee, with the variance, the likely cause and a suggested correction.

FIVE STEPS MAPPED

01 DECIDE	02 GROUND	03 CALIBRATE	04 PROVE	05 OPERATE

DECIDE	Every fee variant the firm uses: flat, tiered, breakpoint, household.
GROUND	Agreements hold the rates. Custodian feeds hold the AUM.
CALIBRATE	Last quarter's billing, against corrections humans already caught.
PROVE	The controller approves each flag before any adjustment.
OPERATE	Each billing cycle, split clean from exception.

RESULT

~~3 to 5 days per cycle~~ — A same-morning variance report

Vendor Invoice Validation

Every invoice checked against the purchase order and the policy before the payment leaves.

WHO FEELS IT · CONTROLLER / VP FINANCE

WHAT IT TAKES TODAY

15–30 MINUTES PER INVOICE

An AP analyst matches each invoice to its purchase order and receipt, re-checks totals, freight and tax, verifies the approval threshold, and chases sign-offs by email.

INDUSTRY ESTIMATE

A mid-size firm processes thousands of invoices a quarter, each carrying its own threshold, variance and approval rules.

WHAT THE AGENT DOES

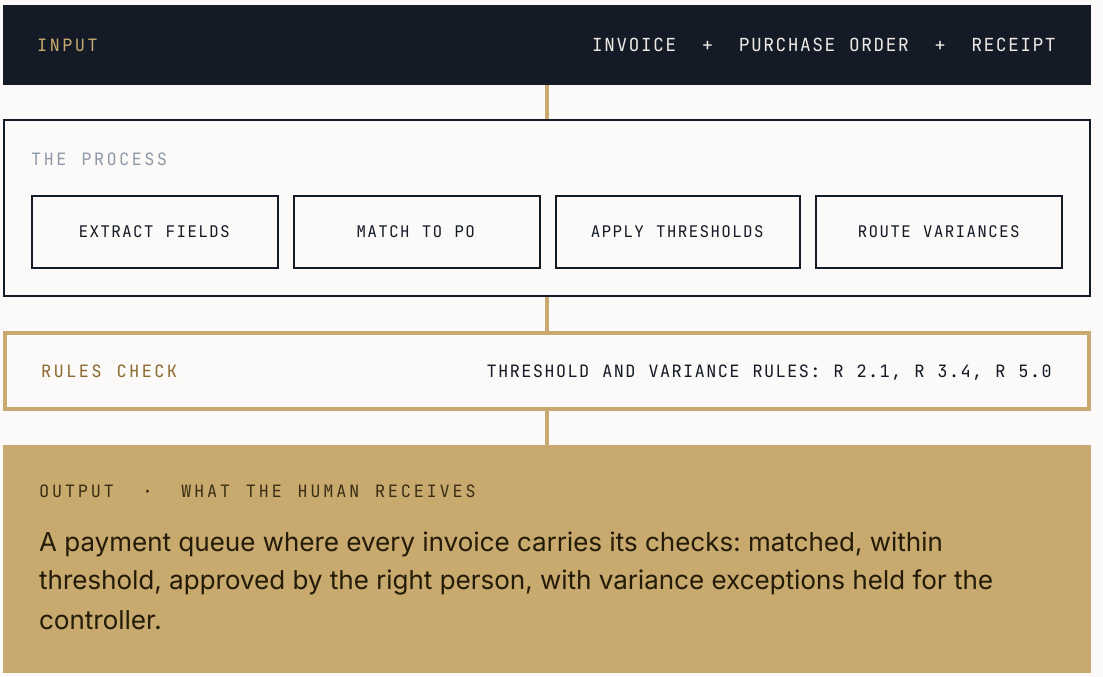
01 Extracts vendor, amounts, line items, tax and payment terms from the incoming invoice.

02 Matches the invoice to its purchase order and goods receipt, line by line.

03 Applies the firm's thresholds: auto-approve under $5K on an exact match, manager sign-off to $10K, VP Finance above.

04 Routes any variance above 2% from the PO to the controller, with the delta and the likely cause.

05 Queues approved payments and logs every check, threshold and sign-off as it goes.



OUTPUT · WHAT THE HUMAN RECEIVES

A payment queue where every invoice carries its checks: matched, within threshold, approved by the right person, with variance exceptions held for the controller.

FIVE STEPS MAPPED

01 DECIDE	02 GROUND	03 CALIBRATE	04 PROVE	05 OPERATE

DECIDE	Thresholds, variance limits and who approves at each tier, signed by VP Finance.
GROUND	The AP system is truth for amounts. The ERP holds the purchase orders.
CALIBRATE	A quarter of historical invoices, including the ones corrected by hand.
PROVE	Every payment checked against R 2.1, R 3.4 and R 5.0 before it queues.
OPERATE	Per invoice, on arrival, each check attached before anything queues.

RESULT

~~15 to 30 minutes per invoice~~ — Exceptions only, everything else queued with proof

Client Onboarding Form Validation and Routing

Engagements start on the day the advisor submits, not three round-trips later.

WHO FEELS IT · OPERATIONS MANAGER / HEAD OF FINANCE

WHAT IT TAKES TODAY

9 in 10 FORMS ARRIVE INCOMPLETE

INDUSTRY ESTIMATE

Finance receives a form with fields missing, emails the advisor for corrections, waits days for a reply, and repeats the cycle two or three times per new engagement.

WHAT THE AGENT DOES

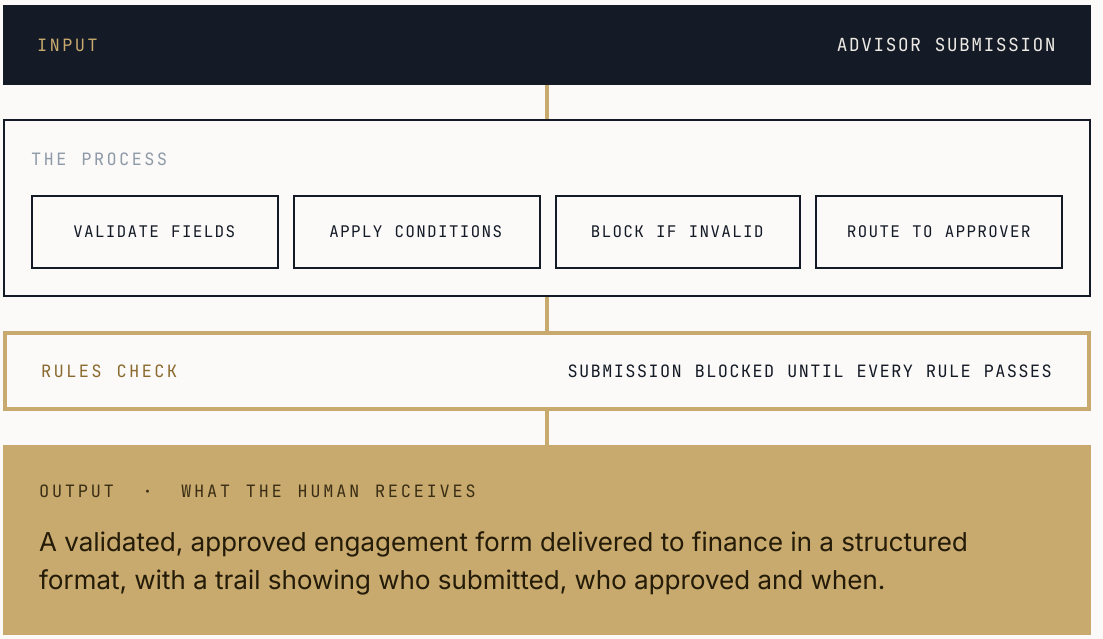
01 Presents a form with mandatory fields, conditional logic and validation as the advisor types.

02 Blocks submission until every rule passes, so incomplete forms never leave the advisor's desk.

03 Routes the completed form to the division leader for a single approve or reject.

04 On reject, returns it to the advisor with specific notes on what needs correcting.

05 On approval, delivers the validated form to finance with every field structured for processing.



FIVE STEPS MAPPED

01 DECIDE	02 GROUND	03 CALIBRATE	04 PROVE	05 OPERATE

DECIDE	Every field: which are mandatory, which are conditional on another answer.
GROUND	Fee schedules and the approval hierarchy come from firm policy.
CALIBRATE	Twenty forms with planted errors against every validation rule.
PROVE	Division leader approval sits inside the flow, not beside it.
OPERATE	Live on every engagement. No form reaches finance unvalidated.

RESULT

~~9 in 10 forms incomplete~~ — No form reaches finance unvalidated

Expense Policy Audit

The $9,400 dinner gets caught before it is reimbursed, not three months later.

WHO FEELS IT · CONTROLLER / CFO

WHAT IT TAKES TODAY

5–10 MINUTES OF REVIEW PER REPORT

INDUSTRY ESTIMATE

A controller reads each report against the policy: per-head caps, spend-class limits, receipt rules, pre-approvals. Most reports are fine. The expensive one looks fine too.

WHAT THE AGENT DOES

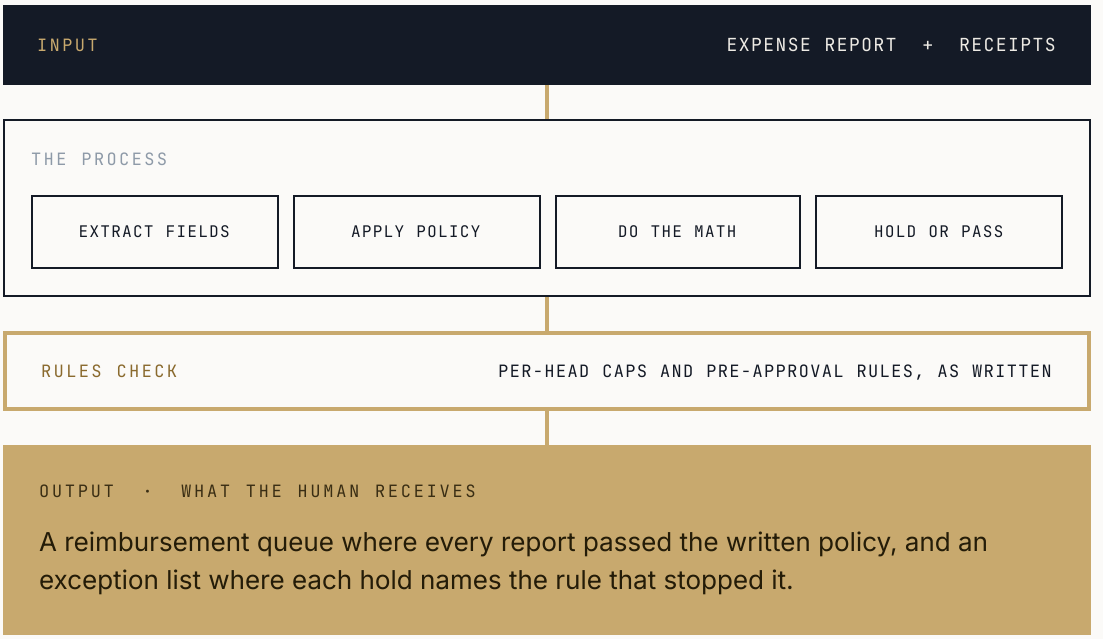
01 Extracts merchant, amount, attendees, spend class and receipt data from each submitted report.

02 Applies the written policy: per-head caps, per-event limits, class rules, pre-approval requirements.

03 Does the math a pattern match skips: $9,400 across two attendees is $4,700 a head against a $150 cap.

04 Flags the violation with the rule named and the delta shown, and holds the reimbursement.

05 Routes clean reports for payment and exceptions to the controller, logging every check.



OUTPUT · WHAT THE HUMAN RECEIVES

A reimbursement queue where every report passed the written policy, and an exception list where each hold names the rule that stopped it.

FIVE STEPS MAPPED

01 DECIDE	02 GROUND	03 CALIBRATE	04 PROVE	05 OPERATE

DECIDE	Every cap and limit in the policy, including the per-head rule nobody had put in a spec.
GROUND	The policy as written is truth. The report is the input, never the rule.
CALIBRATE	Last year’s reports, including the ones the controller caught by eye.
PROVE	Every reimbursement checked against the policy before payment. Pass or hold, rule named.
OPERATE	On every submission, with the trail written as the checks run.

RESULT

~~A controller reading every report by eye~~ — Exceptions only, each with the rule named

Wire and ACH Validation

Every transfer checked against standing instructions and approval limits before the money moves.

WHO FEELS IT · OPERATIONS MANAGER / WIRE DESK

WHAT IT TAKES TODAY

15–20 MINUTES PER REQUEST

INDUSTRY ESTIMATE

A wire desk analyst pulls the client's standing instructions, verifies the bank details match, checks the amount against approval thresholds, routes to signers, and documents each step for audit.

WHAT THE AGENT DOES

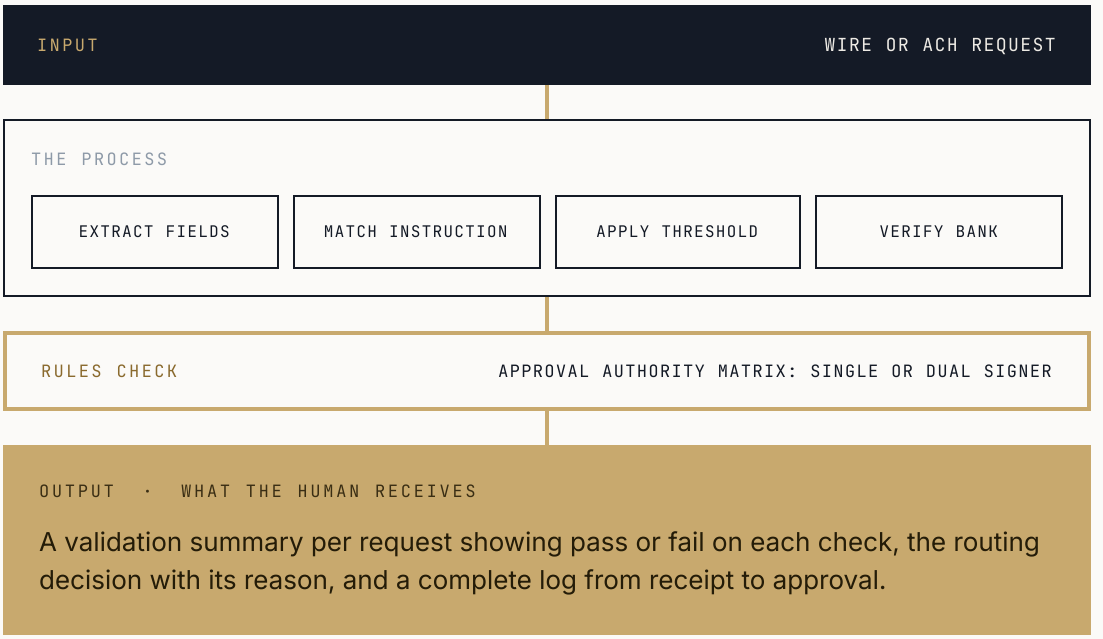
01 Extracts the sending account, receiving bank, beneficiary account, amount and stated purpose.

02 Checks whether a matching standing instruction exists and whether every detail agrees with it.

03 Applies the threshold rules: single signer below the firm limit, dual signer above it.

04 Verifies bank details against known-good records, including routing validation and format checks.

05 Routes to the approvers with a summary of what passed and what needs judgment, logging every step.



OUTPUT · WHAT THE HUMAN RECEIVES

A validation summary per request showing pass or fail on each check, the routing decision with its reason, and a complete log from receipt to approval.

FIVE STEPS MAPPED

01 DECIDE	02 GROUND	03 CALIBRATE	04 PROVE	05 OPERATE

DECIDE	Threshold tiers, approval authority, fields that must match exactly.
GROUND	Standing instructions are truth. Firm policy sets the thresholds.
CALIBRATE	Forty historical requests, including ten that were rejected.
PROVE	No wire executes without approval. The agent validates, never authorizes.
OPERATE	Real time, per request, with the log written as it goes.

RESULT

~~15 to 20 minutes per request~~ — A validated summary waiting at the desk

Policy Bound: Setup and Calendar

Nothing about a bound policy depends on someone remembering to key it in.

WHO FEELS IT · ACCOUNT MANAGER / AGENCY OPERATIONS

WHAT IT TAKES TODAY

30+ MINUTES PER BINDING

INDUSTRY ESTIMATE

An account manager reads the declaration page, keys carrier, policy number, insured, limits, dates and premium into the book-of-business system, sets a reminder, drafts a welcome note and files the document.

WHAT THE AGENT DOES

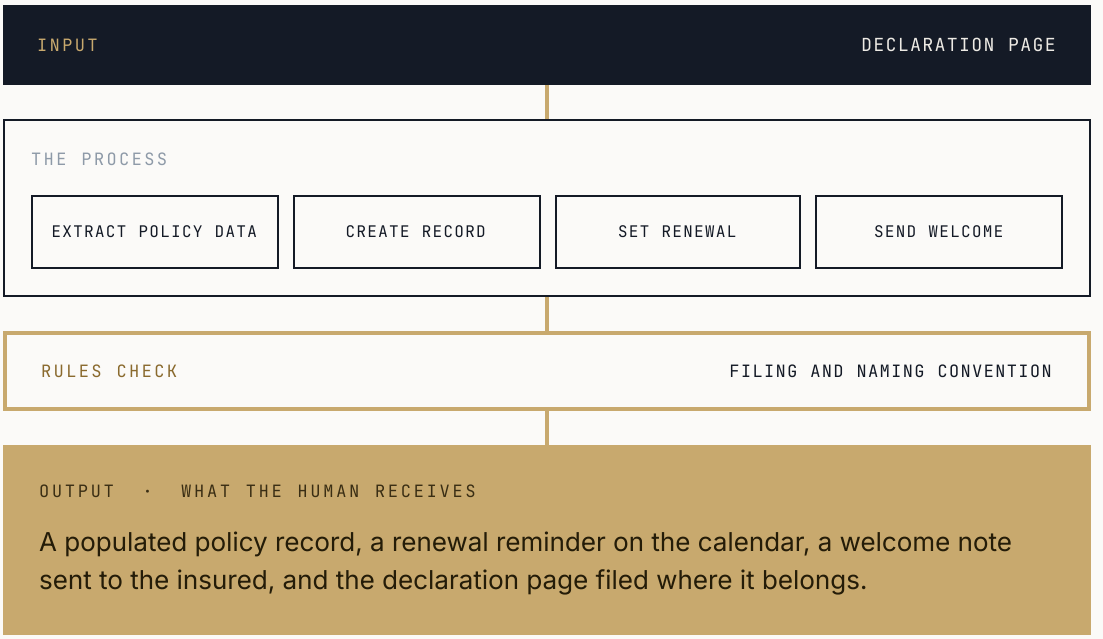
01 Extracts insured name, carrier, policy number, coverage, limits, deductibles, dates and premium from the declaration page.

02 Creates the policy record in the book-of-business system with every field populated.

03 Sets a renewal reminder ninety days before expiry, assigned to the responsible account manager.

04 Sends a welcome communication to the insured confirming coverage, with the key details summarized.

05 Files the declaration page in the correct client folder under the firm's naming convention.



OUTPUT · WHAT THE HUMAN RECEIVES

A populated policy record, a renewal reminder on the calendar, a welcome note sent to the insured, and the declaration page filed where it belongs.

FIVE STEPS MAPPED

01 DECIDE	02 GROUND	03 CALIBRATE	04 PROVE	05 OPERATE

DECIDE	Every declaration page field, plus filing and naming conventions.
GROUND	The declaration page is truth. The management system is the target.
CALIBRATE	Thirty historical pages across five carrier formats.
PROVE	The account manager confirms or corrects what was entered.
OPERATE	On every binding notification, with no entry step left.

RESULT

~~30 or more minutes per binding~~ — A record, a reminder and a filed page

KYC Collection and Verification

Compliance spends its time reading documents rather than chasing them.

WHO FEELS IT · COMPLIANCE OFFICER / AML ANALYST

WHAT IT TAKES TODAY

3–6 WEEKS PER CAMPAIGN

An analyst exports the clients due, sends individual requests, tracks replies in a spreadsheet, sends follow-ups by hand, and reviews each document for validity when it finally arrives.

INDUSTRY ESTIMATE

A refresh campaign is estimated to need five or more manual follow-ups per client.

WHAT THE AGENT DOES

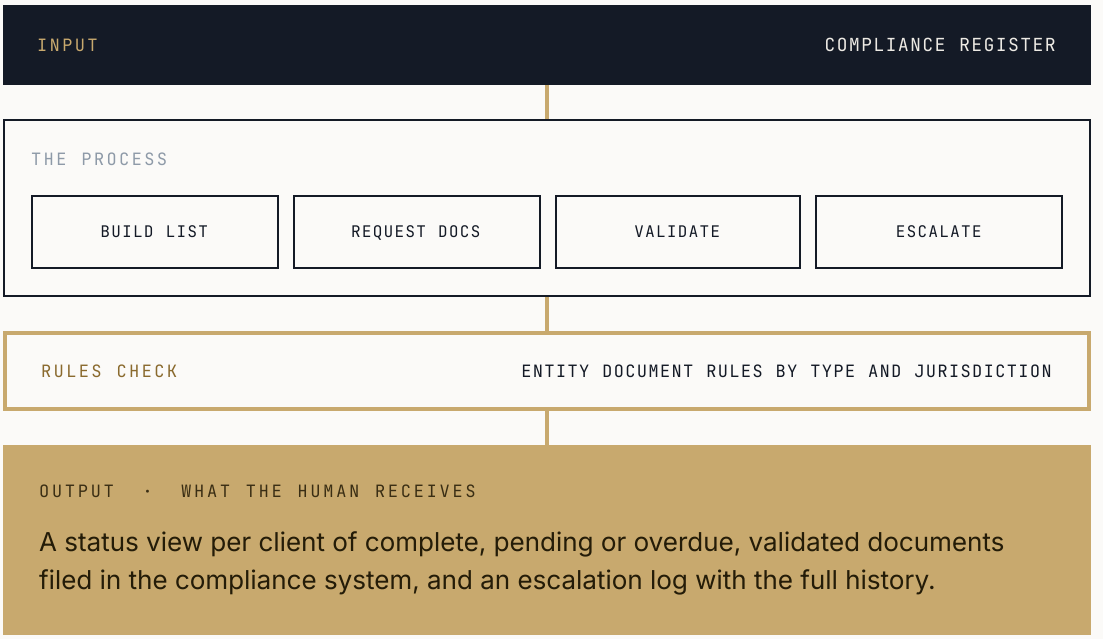
01 Monitors expiry dates and builds a collection list sixty days before each refresh deadline.

02 Sends requests naming exactly which documents are needed for that entity type.

03 Validates what arrives: identity not expired, address proof inside ninety days, entity papers matching the registered name.

04 Updates the compliance tracker with status, receipt date and validation result per client.

05 Escalates non-responders on schedule: reminder at day seven, notice at fourteen, relationship manager at twenty-one.



OUTPUT · WHAT THE HUMAN RECEIVES

A status view per client of complete, pending or overdue, validated documents filed in the compliance system, and an escalation log with the full history.

FIVE STEPS MAPPED

01 DECIDE	02 GROUND	03 CALIBRATE	04 PROVE	05 OPERATE

DECIDE	Document requirements per entity type and jurisdiction.
GROUND	The compliance register defines who is due and when.
CALIBRATE	A mock campaign of fifty records with expired and mismatched papers.
PROVE	Compliance reviews any document that fails validation.
OPERATE	Continuous. Collection at sixty days, escalation on schedule.

RESULT

~~3 to 6 weeks per campaign~~ — A live status view, always current

Claims Intake and Coverage Verification

Coverage confirmed and reserves set before the claim reaches an adjuster, not after.

WHO FEELS IT · CLAIMS MANAGER / TPA OPERATIONS

WHAT IT TAKES TODAY

25–40 MINUTES PER NOTICE

INDUSTRY ESTIMATE

A claims handler reads the first notice of loss, pulls the policy, confirms coverage was in force on the date of loss, checks limits, deductibles and exclusions, and keys the claim in before anyone can be assigned.

WHAT THE AGENT DOES

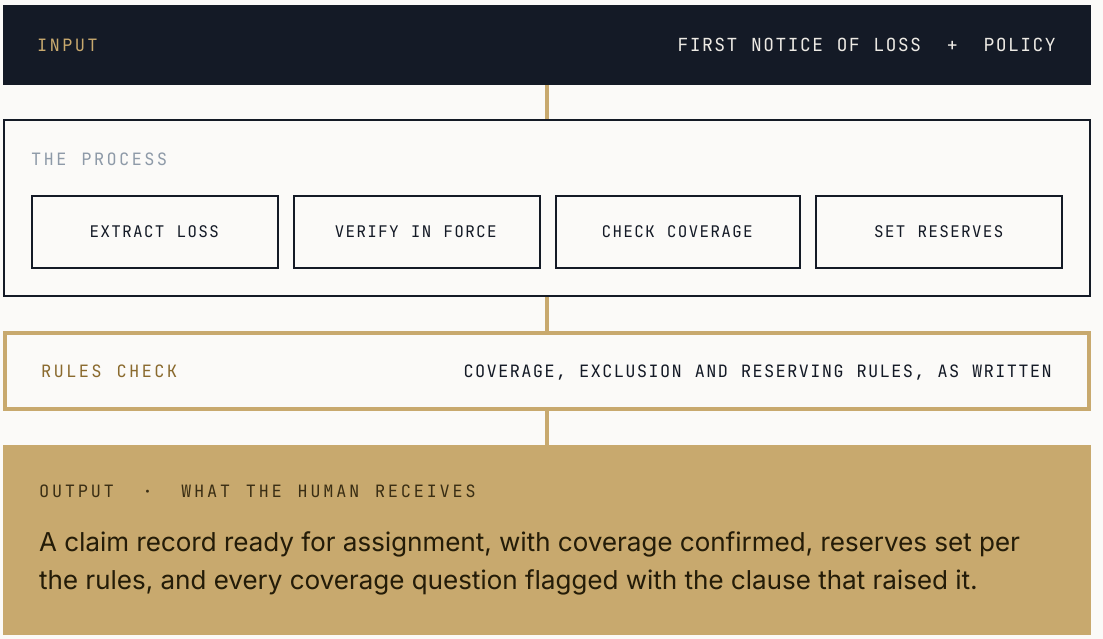
01 Extracts date of loss, cause, parties and damage details from the notice, whatever channel it arrives on.

02 Pulls the policy and confirms it was in force on the date of loss, including lapses and endorsements.

03 Checks the loss against coverage, limits, deductibles and exclusions, as written in the policy.

04 Creates the claim record with reserves set per the firm's reserving rules, and flags coverage questions.

05 Routes clean claims to assignment and exceptions to a senior adjuster, logging every check.



OUTPUT · WHAT THE HUMAN RECEIVES

A claim record ready for assignment, with coverage confirmed, reserves set per the rules, and every coverage question flagged with the clause that raised it.

FIVE STEPS MAPPED

01 DECIDE	02 GROUND	03 CALIBRATE	04 PROVE	05 OPERATE

DECIDE	What counts as covered, excluded and escalation-worthy, signed by claims leadership.
GROUND	The policy admin system is truth for coverage. The notice is input, never authority.
CALIBRATE	Sixty historical claims, including denials that were later overturned.
PROVE	No coverage position reaches an insured without the clause named. Pass or hold.
OPERATE	On every notice, any channel, with the file built as it goes.

RESULT

~~25 to 40 minutes per notice~~ — Assignment-ready claims, exceptions flagged

Loan File Completeness and Policy Check

Underwriters underwrite. The file arrives complete, calculated and checked against policy.

WHO FEELS IT · CHIEF LENDING OFFICER / HEAD OF UNDERWRITING

WHAT IT TAKES TODAY

2–3 DAYS OF CHASE PER FILE

INDUSTRY ESTIMATE

A loan processor checks each file for required documents, recalculates ratios, verifies the file against lending policy, and emails the member or the officer for whatever is missing, file by file.

WHAT THE AGENT DOES

01 Checks the application file against the required-document list for that loan type.

02 Extracts income, obligations, collateral value and score data from the file's documents.

03 Recalculates DTI and LTV and checks them against the lending policy's thresholds.

04 Flags policy exceptions with the rule named, and requests missing items in one pass, not five.

05 Delivers underwriting-ready files with the checks and the trail attached.

INPUT — LOAN APPLICATION FILE

THE PROCESS

- CHECK DOCS
- EXTRACT DATA
- RECALCULATE RATIOS
- FLAG EXCEPTIONS

RULES CHECK — LENDING POLICY: DTI, LTV AND EXCEPTION THRESHOLDS

OUTPUT · WHAT THE HUMAN RECEIVES

An underwriting-ready file: documents complete, ratios calculated and checked against policy, exceptions named, and one request list for anything missing.

FIVE STEPS MAPPED

01 DECIDE	02 GROUND	03 CALIBRATE	04 PROVE	05 OPERATE

DECIDE	Required documents and thresholds per loan type, signed by the credit committee.
GROUND	The core system and the file are read live. The policy as written sets the thresholds.
CALIBRATE	Fifty funded files and ten declines, against what underwriting actually did.
PROVE	No file reaches underwriting with an unchecked ratio. Exceptions carry the rule.
OPERATE	On every application, with completeness enforced before review time is spent.

RESULT

~~2 to 3 days of chase per file~~ — Underwriting-ready files, first pass

09.2 / INTEGRATION MAP

A connection, not a data project.

Step two asks which systems hold the truth. These are the ones already connected, which means the answer is a configuration rather than a build.

CUSTODIANS	Schwab · Fidelity · Pershing · Wells Fargo
PORTFOLIO SYSTEMS	Orion · Black Diamond · Tamarac
CRM	Wealthbox · Salesforce · Redtail
COMPLIANCE	RegEd · ComplySci
DOCUMENT	DocuSign · ShareFile · Google Drive
COMMUNICATION	Outlook · Gmail
BILLING	Orion Billing · BillFin

PLUS SPREADSHEETS, SFTP DROPS AND ANY DOCUMENTED API

40+

SYSTEMS CONNECTED

PitCrew Factory, the platform in chapter ten, reads from the stack you already run. Nothing is migrated, and no system is asked to become the system of record unless you name it as one.

Different operations. The same five steps.

	PROCESS	WHO FEELS IT	TODAY
01	IPS Drift Compliance Check	CHIEF COMPLIANCE OFFICER	40+
02	Fee Billing Validation	CONTROLLER	3–5
03	Vendor Invoice Validation	CONTROLLER / VP FINANCE	15–30
04	Client Onboarding Form Validation and Routing	OPERATIONS MANAGER	9 in 10
05	Expense Policy Audit	CONTROLLER / CFO	5–10
06	Wire and ACH Validation	OPERATIONS MANAGER	15–20
07	Policy Bound: Setup and Calendar	ACCOUNT MANAGER	30+
08	KYC Collection and Verification	COMPLIANCE OFFICER	3–6
09	Claims Intake and Coverage Verification	CLAIMS MANAGER	25–40
10	Loan File Completeness and Policy Check	CHIEF LENDING OFFICER	2–3

WHAT THE TEN HAVE IN COMMON

Every one of them was already a written-down process somewhere in the firm. None needed a smarter model. Each needed the firm's judgment written down and the output made checkable.

Every one keeps a human on the approval. The agent does the reading, the matching and the logging. The firm keeps the decision and the signature.

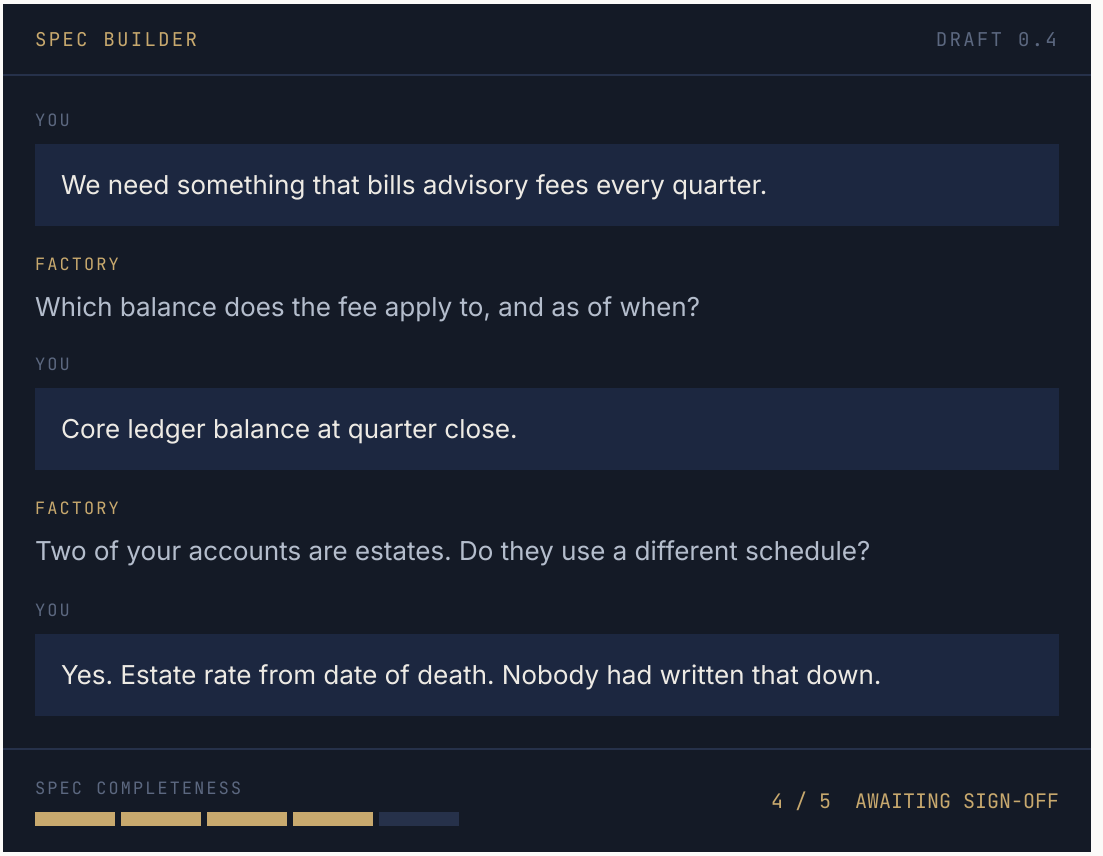
Chapter eleven gives five tests for choosing among them. Start where the cost is already visible.

Factory

Your operational knowledge is the asset. It lives in people's heads, in policies, in exception lists nobody has collected. This is how it becomes a system that runs, gets tested, and produces proof on every pass.

From a sentence to a running agent.

You describe a process in plain language. Factory's Spec Builder asks the questions that make the spec complete: what triggers it, what rules apply, what exceptions exist, who approves. When the spec is signed off, Factory builds.



The questions are the point. Most of what a spec needs is already known inside the firm. It has simply never been asked for in one sitting.

Every agent walks the same line.

No step is optional. A failure at any stage blocks the one after it.

01 **Decide** — SPEC SIGNED

Spec Builder captures the process and the owner signs it.

02 **Ground** — 6 CONNECTED

Connect to your systems. 40+ pre-built integrations to custodians, CRMs and portfolio systems.

03 **Calibrate** — 47 SCENARIOS

Run against your scenarios. Failures block deployment.

04 **Prove** — RULES HOLD

AWS Automated Reasoning formally checks that your rules hold. Pass or fail.

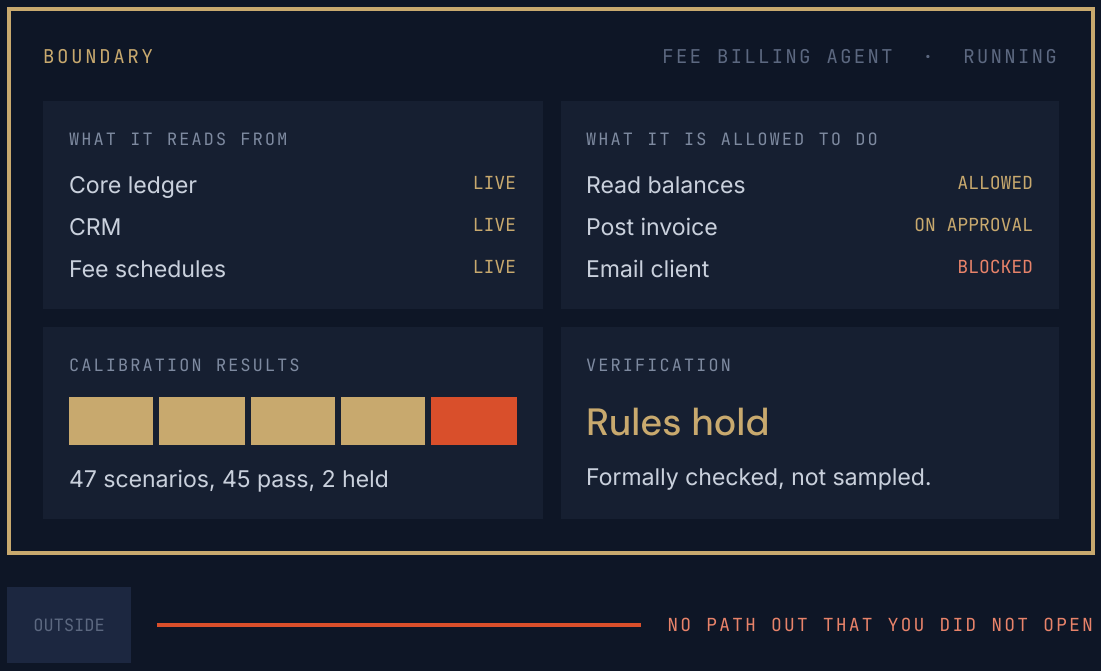
05 **Operate** — AWAITING APPROVAL

Your team approves. Proof from the first action.

Fee billing agent, build 0.4 — 4 OF 5 STAGES CLEARED

Powerful inside the boundary. Nothing outside it.

Every agent runs inside a boundary you set. It can reach only the systems and data you connect, and nothing outside them. Anything else is blocked unless you allow it.



WHAT THE BOUNDARY IS FOR

The agent can be capable without being unsupervised. Capability is granted, not assumed.

WHAT IT PRODUCES

A list of exactly what the agent could reach, which is the first thing any reviewer asks for.

Not a black box. A record you can open, with every version kept.

When a regulation, a custodian or a policy changes, you update the spec and Factory rebuilds. Same verification. Same trail.

Quarterly Fee Billing

v2.3

NEEDS

Core ledger balances
Fee schedules
Household groupings

PRODUCES

Invoice per account
Exception queue
Sealed proof package

PROCESS STEPS, TYPED AND VISIBLE

READ | MATCH | CALCULATE | PROVE | APPROVE | POST

RUN HISTORY

30 SEP 26	1,184 accounts, 3 exceptions	PASS
30 JUN 26	1,171 accounts, 1 exception	PASS
31 MAR 26	1,166 accounts, 8 held on R 5.0	HELD

Everything a reviewer would ask for is a property of the artifact, not a document someone has to write afterwards.

v2.2 TO v2.3

Per-head cap added after the March holds. Spec updated, agent rebuilt.

RE-VERIFIED

All 47 scenarios re-run against the new rule set before the June cycle.

You do not have to start from zero.

Adopt

DAYS

Take a pre-built agent from the catalog of 60+ Skills and configure it to your rules. The spec, the tests and the verification layer already exist. You supply the judgment that is specific to your firm.

Adapt

A WEEK

Start from a working spec and change what differs. Different data sources, different rules, different approval path. The shape of the process is already proven.

Build

YOUR PACE

Something new, from scratch, with the Spec Builder. Your process, your spec, your timeline. The five steps are the same in every case.

Most firms start by adopting one agent for a process that already hurts, and learn the discipline on something small enough to be safe. Any model runs underneath, open source or frontier, so your costs follow your own choices rather than one vendor’s pricing.

The methodology is free. The platform is what happens when you stop enforcing it by hand.

SPEC BUILDER	Step one becomes a conversation instead of a blank document.
INTEGRATIONS	Step two becomes a connection, not a data project.
SCENARIO SUITE	Step three blocks deployment instead of producing a report nobody reads.
AUTOMATED REASONING	Step four is a formal check, not a second opinion. Powered by AWS Automated Reasoning.
BOUNDARY AND PROOF	Step five produces the proof while the work is happening.

You can keep funding investments that demo well.

Or you can build so the answer still holds the day someone asks how you got it.

CHAPTER ELEVEN

Put It Into Practice

Three actions any firm can take right now. No tools required. Takes a morning.

ACTION 01 Write down what “correct” looks like

Choose the process that causes the most pain: fee billing, onboarding, reconciliation. Sit with whoever owns it. Write down what triggers it, what steps happen, what rules apply, what exceptions exist, and what the output should look like. Get sign-off. You now have a spec. Most firms have never done this for their most critical processes.

ACTION 02 Find where the data actually lives

For that same process, list every system that holds the data. Custodian, CRM, portfolio system, shared drive, email, spreadsheet. Then ask which one is the source of truth when they disagree. If nobody can answer, that is your grounding problem.

ACTION 03 Find your five hardest scenarios

Pull five real examples from the past year where the process was tricky: the exceptions, the unusual accounts, the ones that took extra time. These become your first calibration set.

None of this costs anything. It will tell you whether your firm is ready to build on its own judgment, or is still running on assumptions.

Five tests for the first operation to govern.

Before anyone builds anything, decide where. The right first operation passes all five.

01 **The error is denominated in dollars.**

A wrong output becomes a refund, a clawback, a penalty, an overpayment. If an error costs nothing, govern something else first.

02 **The rules already exist.**

Written down, or writable in one sitting by the person who owns the process. If nobody can state what correct means, that is a decision problem, not an AI problem.

03 **The volume is high enough to matter.**

Daily or per-transaction, not quarterly. The discipline's return compounds on frequency.

04 **Someone senior already owns it.**

A signature is available for the spec. Processes without an owner stall at step one.

05 **Proof is already being demanded.**

A board, an auditor, a regulator or a client is already asking how you know. The record pays for itself from the first run.

Fee billing, vendor payments, claims, onboarding, reconciliation: most firms have three candidates that pass all five. Pick the one that already hurts. The discipline is easier to learn on a process you can describe from memory.

One process. One page. One signature.

Italic lines are examples, not answers. Write over them.

PROCESS

e.g., Quarterly fee billing reconciliation

WHAT TRIGGERS IT

e.g., Signed advisory agreement received from client

WHAT RULES APPLY

e.g., Billable balance from core ledger at quarter close

WHAT EXCEPTIONS EXIST

e.g., Estate accounts use the estate rate from date of death

WHAT “CORRECT” LOOKS LIKE

e.g., Invoice matches the calculated fee within $0.01

SIGNED OFF BY

DATE

WORKSHEET 02 GROUND

Where the data actually lives.

SYSTEM	WHAT IT HOLDS	TRUTH?
e.g., Core ledger	*e.g., Balances and positions*	☐
		☐
		☐
		☐
		☐
		☐

Exactly one box should be ticked. If two are, that is the argument to have before anything is built.

WORKSHEET 03 CALIBRATE

Your five hardest scenarios.

01 *e.g., Trust within a trust*

02 *e.g., Deceased beneficiary mid-quarter*

03

04

05

Run your own numbers.

The Return, from chapter two, with your figures in it. Italic lines are examples, not answers.

ERRORS REACHING CLIENTS, PER YEAR $

e.g., fee or billing errors found at the annual review

SENIOR HOURS RECONSTRUCTING ANSWERS, PER YEAR $

e.g., 40 hours per "how do you know?" × requests × loaded rate

THE LAST KEY DEPARTURE $

e.g., six months of rebuilt judgment, errors in the gap

HEADCOUNT ADDED FOR GROWTH, LAST YEAR $

e.g., ops hires that scale 1:1 with accounts won × loaded cost

TOTAL, CARRIED EVERY YEAR $

Compare the total to the cost of adopting the discipline on one operation. That is the business case, in your numbers, on one page.

The questions your board will ask.

You will be asked these. The book’s answers, in one place.

“How do we know the output is right?”

STEP FOUR · PAGE 38

Because a separate system checked it against your written rules and said pass, with the rule named. Confidence scores measure the model; they say nothing about this answer.

“What is the P&L impact?”

THE RETURN · PAGE 18

Count what you carry now: errors reaching clients, senior hours reconstructing answers, headcount that scales with growth. Worksheet 04 turns it into your number.

“Is our data safe?”

THE BOUNDARY · PAGE 73

The agent reaches only the systems you connect, inside a boundary you set. There is no path out you did not open.

“Our data is a mess. Fix that first?”

STEP TWO · PAGE 30

No. Name one source of truth per field that matters and connect it live. That is a decision, not a data project.

“Build or buy?”

CHAPTER TEN · PAGE 70

The discipline is free and vendor-agnostic either way. Buy the enforcement when running the five steps by hand stops scaling.

“What happens to our team?”

STEP FIVE · PAGE 42

They stop typing and start approving. The agent cannot sign off on its own work; your people keep the decision and the signature.

CHAPTER TWELVE

Where To Go From Here

The discipline is published and vendor-agnostic. You can follow it with any combination of tools.

If you want a platform where all five steps are built in, with spec templates, integrations, calibration scenarios, verification and proof produced automatically, that is Factory.

The methodology

All five steps, free and open.

gopitcrew.com/the-discipline

Factory

The platform with the discipline built in.

gopitcrew.com/factory

BUILD IT YOURSELF

On Factory, at your own pace, with the five steps enforced by the platform.

BUILD IT WITH US

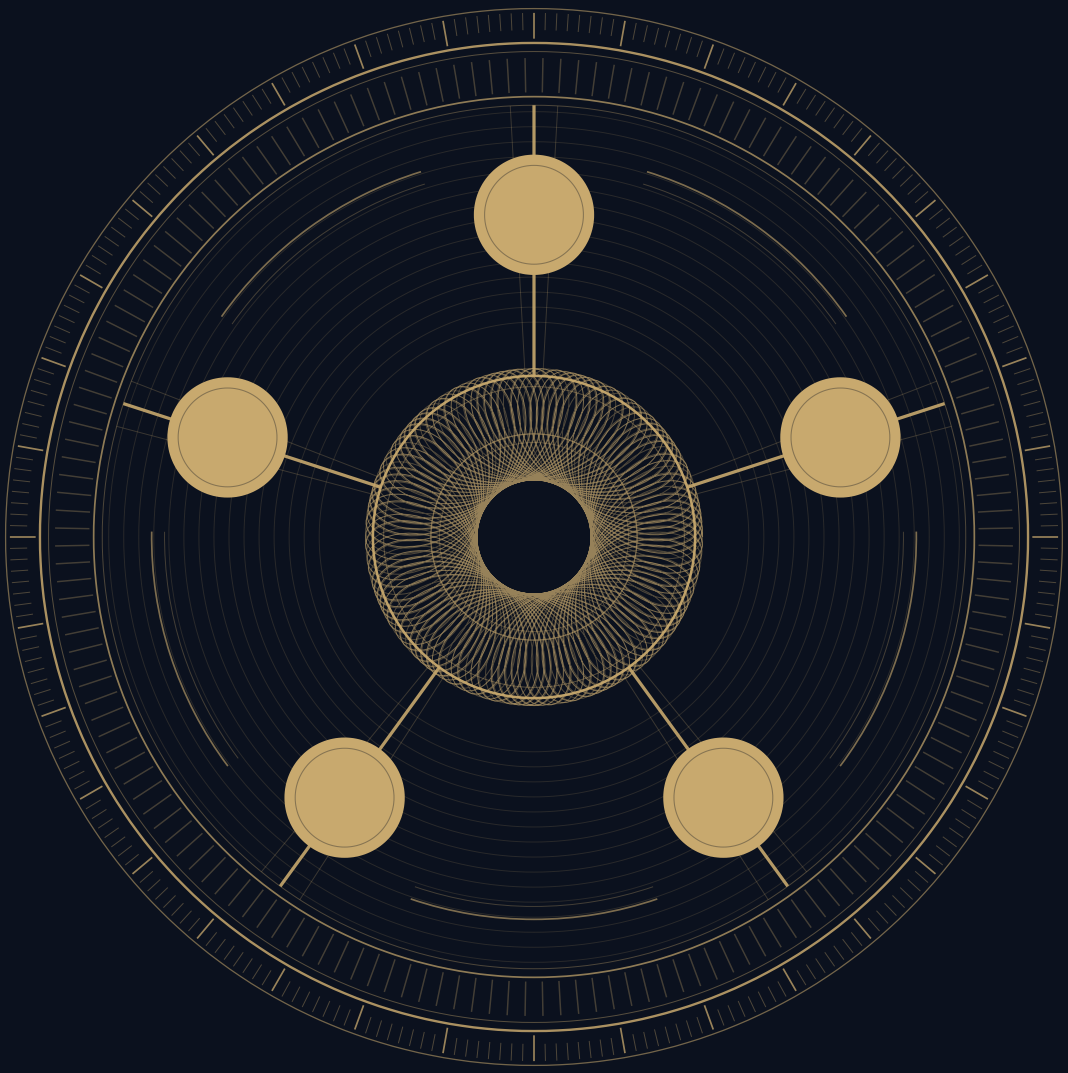
A Forward Deployed Expert builds it with your team in 15 working days.

CONTACT

sales@gopitcrew.com

ALL FIVE THROWN

LOCKED



PITCREW

The agentic AI platform and team for firms where the cost of being wrong is real.

GOPITCREW.COM