The Procurement Function That Never Sleeps

Always-on agents and the rise of autonomous source-to-pay

Saquib Jawed

2026 EDITION

Contents

FRONT MATTER

Foreword

Executive Summary

Key Figures

PART I — WHY NOW

01 The Attention Gap

02 What “Always-On” Actually Means

PART II — HOW IT WORKS

03 Mandates, Not Prompts

04 Inside the Machine

05 Twenty-Four Hours in an Autonomous Procurement Function

PART III — TRUST AND ECONOMICS

06 Earned Autonomy

07 The Economics of Never Sleeping

08 The Supervisory Procurement Organization

PART IV — WHAT COMES NEXT

09 The Road to 2030

BACK MATTER

Appendix: Readiness Diagnostic, Vendor Questions, Glossary

Sources and Notes

FOREWORD

Every procurement system I have shipped had the same flaw. It waited for someone to log in.

For twenty-five years we digitized procurement. Paper requisitions became web forms, filing cabinets became contract repositories, and fax approvals became workflow queues. The work got faster. The model did not change: a person notices something, opens a system and tells it what to do.

Copilots made that person sharper once they arrived. They did not change *when* the person arrives, and in source-to-pay that is where the money goes. Renewals lapse overnight. Commodity prices fall on a weekend. A supplier's credit rating slips while the category manager is on leave. None of these events is hard to detect. Each one is missed because the organization can only pay attention during business hours.

This book argues that the next platform shift in procurement is not smarter software. It is software that keeps watch: agents that hold standing intent, sense change continuously, act within clear limits and involve people only when judgment is required.

It is written for chief procurement officers, finance leaders and the transformation teams who will govern these agents. Where I use numbers, they are sourced. Where I use scenarios, they are labeled as illustrative. The aim is a practical point of view that a leadership team can debate, adopt and measure.

A note on vantage point: I build procurement software for a living. This book names no vendor, my employer included, and its frameworks are meant to hold whichever platform you choose.

Saquib Jawed

Bengaluru, November 2026

EXECUTIVE SUMMARY

Seven claims

1 **Attention, not process, is procurement's binding constraint.**
Most value is lost in the interval between a signal appearing and a person noticing it.

2 **Always-on is an architecture, not a copilot left running.**
Five properties distinguish genuine agents from relabeled assistants.

3 **The interface moves from prompts to mandates.**
People author standing intent; agents hold it around the clock.

4 **Reliability is an infrastructure problem before it is a model problem.**
Sensing, a shared world model, tiered reasoning and reversible actions come first.

5 **Autonomy is earned per decision, not switched on per module.**
Track record promotes an agent; drift demotes it.

6 **The cost model shifts from labor to compute and attention.**
Leaders manage what it costs to wake an agent, and what it costs to wake a person.

7 **Procurement teams become supervisors of autonomous capacity.**
By the end of the decade, they will negotiate with suppliers' agents at machine speed.

KEY FIGURES

The numbers behind the argument

90%

of B2B buying will be intermediated by AI agents by 2028, channeling more than $15 trillion of spend, according to a 2025 Gartner forecast.

Gartner, 2025

40%

of enterprise applications would feature task-specific agents by the end of 2026, up from under 5% in 2025, Gartner forecast in August 2025.

Gartner, August 2025

15%

of day-to-day work decisions forecast to be made autonomously by agentic AI by 2028, from zero in 2024.

Gartner, June 2025

40%+

of agentic AI projects expected to be canceled by the end of 2027, chiefly on cost, unclear value or weak risk controls.

Gartner, June 2025

+8%

projected rise in procurement workload in 2026, while headcount and operating budgets decline.

The Hackett Group, 2026

8.6%

of contract value lost on average to poor contract management; the best performers lose around 3%.

World Commerce & Contracting, 2025

$9.90 vs $2.67

average versus best-in-class cost to process a single invoice.

Ardent Partners, 2026

37%

annualized return forgone each time a 2/10 net 30 early-payment discount is missed.

Calculation: 2 ÷ 98 × 365 ÷ 20

CHAPTER 01

The Attention Gap

The binding constraint in procurement is not process, data or headcount. It is human attention. Most value is lost in the interval between a signal appearing and a person noticing it.

ATTENTION CAPACITY, NOT OFFICE HOURS

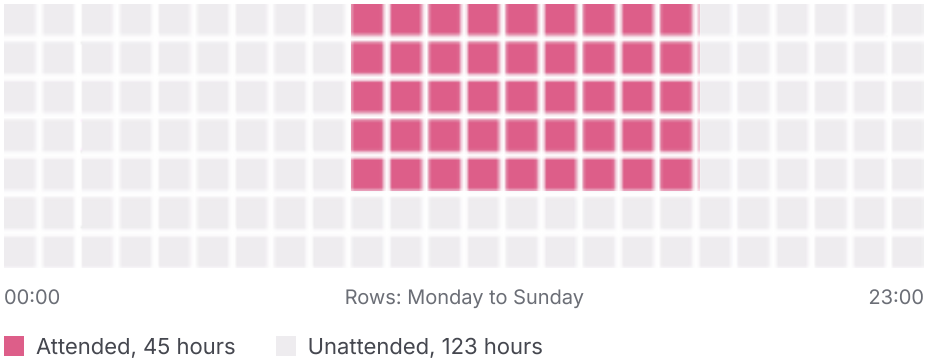
Beneath every procurement suite sits the same architecture: a user logs in, reads a queue and acts. Requisitions, approvals, sourcing events and invoice matching all begin with a human click. The constraint this creates is less about the clock than about capacity: more signals arrive than a team can notice, judge and act on, at any hour.

Leakage has many causes, including poor master data, weak policy and badly drafted contracts, and agents do not fix those. But a material share of the losses on the following pages arises because a known, detectable event was not acted on in time. That share is the subject of this book. Office hours simply make the problem visible: a standard working week covers only 45 of the week's 168 hours.

Exhibit 1.1 One week, by the hour

27%

of the week is staffed in a five-day, nine-to-six model



ANATOMY OF THE GAP

Value is not lost in one place. It drains away across four latencies.

Exhibit 1.2 The four latencies between signal and outcome

01

Signal

Is anyone watching?

Time until the event is visible to someone able to act.

02

Triage

Does it matter?

Time to judge relevance against everything else in the queue.

03

Decision

What do we do?

Time to gather context, consult and approve.

04

Execution

Is it done?

Time to carry out the action across systems and suppliers.

Exhibit 1.3 Four familiar leaks

LEAK	BREAKS AT	WHAT HAPPENS
The silent renewal	Signal	A three-year agreement rolls over at list price because the notice window closed at quarter-end.
The weekend price drop	Triage	An index-linked clause allows repricing, but the alert sits unread among hundreds of others.
The faltering supplier	Decision	Risk flags deterioration for weeks; agreeing on a second source takes longer than the supplier lasts.
The expired discount	Execution	An invoice approved on day six of a ten-day discount window is paid in the next weekly run.

+8%

Procurement workload is projected to rise 8% in 2026 while headcount and budgets decline. There will be more signals and fewer people to notice them. The Hackett Group, 2026.

SIZING THE GAP

Where attention fails, and what it costs

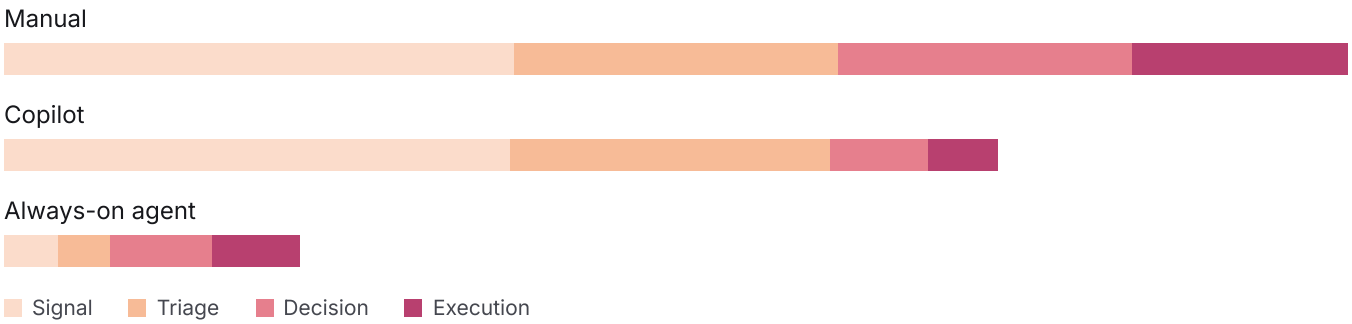
Exhibit 1.4 The Attention Gap Ledger

STAGE	WHERE ATTENTION FAILS	VALUE AT STAKE
Contracts	Scope changes, unapplied price changes and missed obligations after signature	8.6–11% of contract value WorldCC, 2025; WorldCC and Ironclad
Sourcing	Index-linked clauses not exercised; tail spend never competed	1–2% of spend Missed price adjustments; WorldCC and Ironclad
Invoices	Exceptions wait in queues for someone to investigate	19.9% vs 11.8% Average vs best-in-class exception rate; Ardent Partners, 2026
Payments	Discount windows close before approval and payment runs align	≈37% a year Return forgone on a missed 2/10 net 30 discount

WHY COPILOTS DID NOT CLOSE IT

Copilots summarize contracts and draft RFPs in seconds, but they compress only the back half of the gap. Signal and triage still wait for someone to open the tool and ask. A faster person who arrives late is still late.

Exhibit 1.5 Elapsed time from signal to outcome, by operating model (illustrative)



CHAPTER 02

What “Always-On” Actually Means

Always-on is a distinct class of architecture, not a copilot left running in a browser tab. Define it precisely, and it becomes easy to recognize the imitations.

Exhibit 2.1 Five defining properties

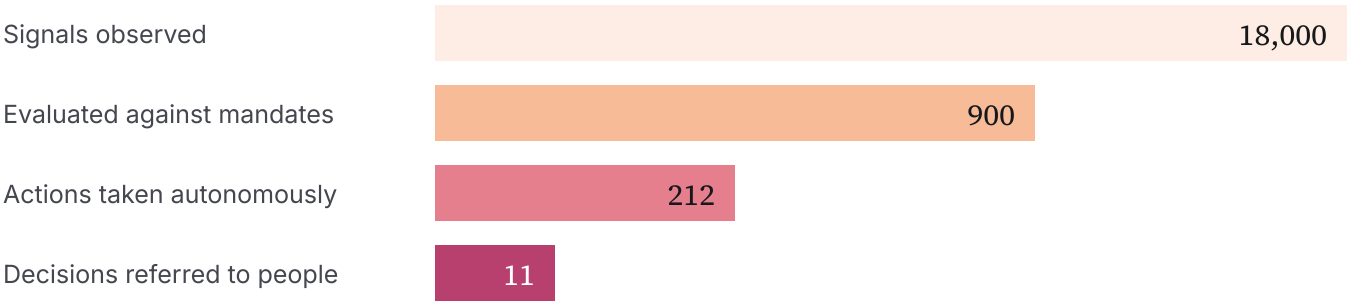
PROPERTY	WHAT IT MEANS	THE TEST
Persistent	Holds state for weeks: open negotiations, pending claims, supplier history.	*What does it remember on day 30?*
Perceptive	Subscribes to internal events and external signals rather than waiting for a prompt.	*What wakes it up?*
Proactive	Starts work on its own when a condition in its mandate is met.	*What did it do unasked?*
Mandated	Acts only within written scope, limits and escalation rules.	*What can it do without asking?*
Accountable	Every action is explainable, logged and, where possible, reversible.	*How do you stop or undo it?*

Remove any one property and the result is useful but different. Persistent without mandated is a liability. Proactive without accountable is a compliance incident. Perceptive without proactive is simply a better dashboard.

ALWAYS-ON IS NOT ALWAYS-ACTING

A well-designed agent is silent most of the time. Its value lies in what it chooses not to do.

Exhibit 2.2 One day at Meridian Components, the company followed in Chapter 5 (illustrative)



Judge an always-on agent by its **interrupt economics**: how much signal it absorbs for every minute of human attention it consumes. An agent that alerts a buyer forty times a day is not autonomous. It is a new inbox.

LESSONS FROM CONSUMER AGENTS

LAYER	WHAT TRANSFERS	WHAT CHANGES IN THE ENTERPRISE
Product	Proactive prompts, standing preferences, a daily digest	A missed prompt costs money, not convenience; every action needs an owner
Infrastructure	Event triggers, schedulers, background execution, memory	ERP write-back must be idempotent, audited and reversible
Harness	Tool use, context assembly, guardrails	Segregation of duties, multiple parties and regulators all apply

THREE GENERATIONS

How always-on agents differ from what came before

Exhibit 2.3 Workflow automation, copilots and always-on agents

	WORKFLOW AUTOMATION	COPILOT	ALWAYS-ON AGENT
Starts work	A rule fires on a form	A person asks	The agent, on a signal
Unit of design	Process flow	Prompt	Mandate
Memory	Transaction record	One session	Weeks of context
Novel cases	Routes to a queue	Reasons, if asked	Reasons, acts or escalates
Human role	Operator	Driver	Governor
Typical failure	Brittle rules	Nobody asked	Acting outside intent

THE IMPOSTER TEST

Gartner estimates that only about 130 of the thousands of vendors claiming agentic AI are building the real thing. Five questions separate them in any demonstration.

1	What wakes it up?	Weak answer: “The user opens it.”
2	What can it do without asking?	Weak answer: everything ends in “recommend”.
3	What does it remember next month?	Weak answer: memory resets each session.
4	How do I stop it and undo its actions?	Weak answer: no stop control, no undo.
5	Who is accountable when it is wrong?	Weak answer: “The AI decided.”

CHAPTER 03

Mandates, Not Prompts

A prompt is an instruction that ends when the answer arrives. A mandate is standing intent that keeps working overnight. The procurement professional's job moves from issuing the first to authoring the second.

Exhibit 3.1 The anatomy of a mandate

01	Objective	The outcome the agent optimizes for
02	Scope	Categories, suppliers, regions and spend bands it may act on
03	Authority	Value and commitment limits for acting alone
04	Constraints	Policy, ESG requirements, preferred suppliers and sanctions
05	Escalation	Conditions that return control to a named person
06	Review	Owner, review cadence and expiry date

Exhibit 3.2 The same need, expressed two ways

A PROMPT

"Find me three quotes for 500 laptops."

A MANDATE

"Keep end-user hardware within 3% of market benchmark. Re-source any line that drifts. Award up to $250k with approved suppliers. Escalate anything else to the IT category lead."

A good mandate is testable: an auditor should be able to say whether any agent action fell inside it. Every mandate also expires, because standing intent without a review date becomes standing risk.

THREE WORKED MANDATES

What standing intent looks like when it is written down

Exhibit 3.3 Sample mandates (thresholds are illustrative)

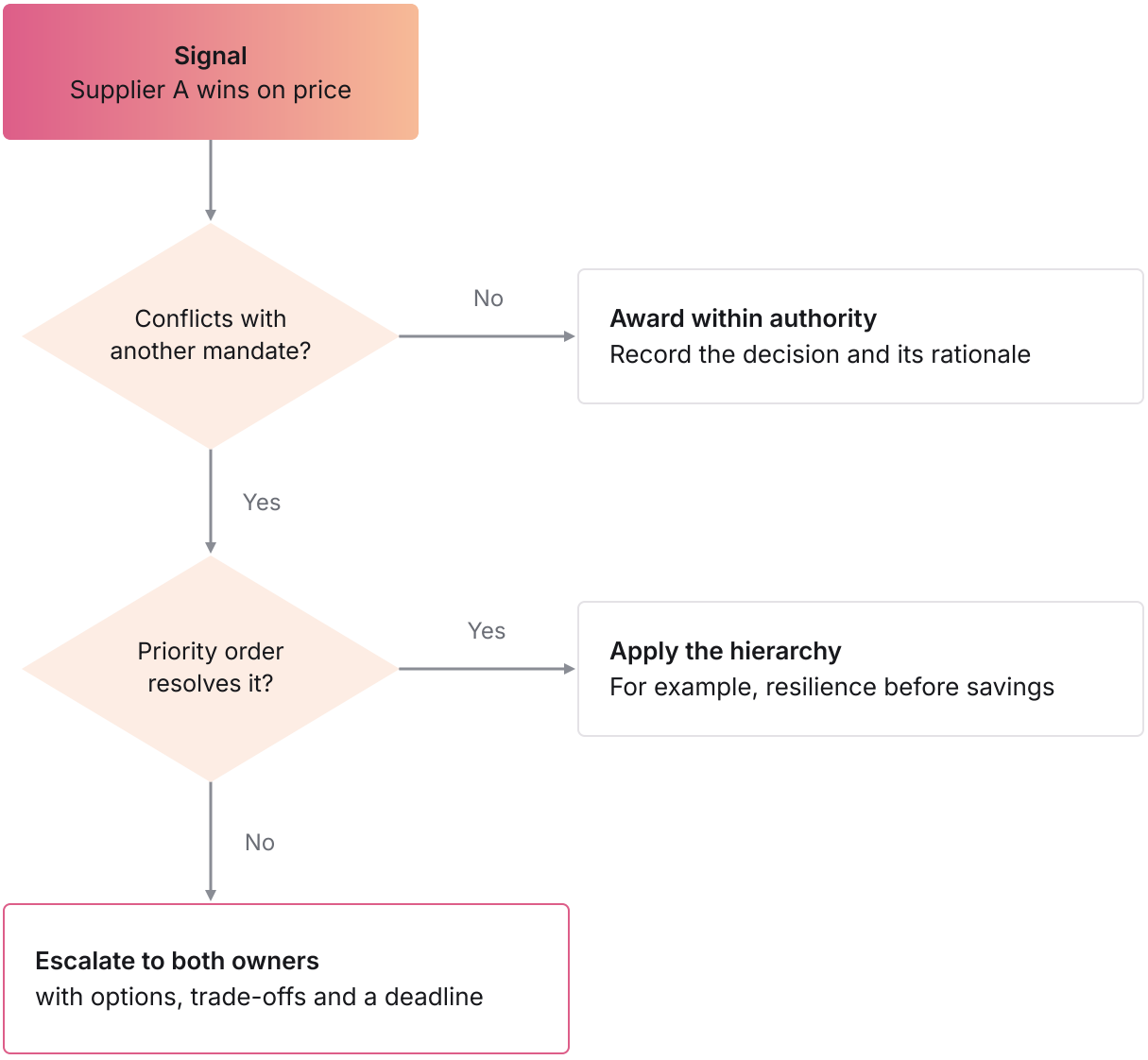
	MND-SRC-014 Tail-spend sourcing	MND-CLM-007 Contract renewal watch	MND-AP-021 Invoice exceptions
OBJECTIVE	Source every indirect request below threshold within 72 hours	No contract renews by default; every renewal is a decision	Resolve price and quantity exceptions within 48 hours
SCOPE	Indirect categories; pre-qualified suppliers; North America	All agreements with auto-renewal or notice clauses	PO-backed invoices across all entities
AUTHORITY	Award up to $50k; three negotiation rounds	Issue non-renewal notices; draft renegotiation briefs	Clear variances within tolerance; request credit notes
CONSTRAINTS	Minimum three bids; no sanctioned entities	Never terminate a sole-source supplier	Never change supplier bank details
ESCALATE WHEN	Single bid; price more than 10% over benchmark	Notice window under 30 days; spend above $1M	Variance beyond tolerance; suspected duplicate
OWNER	Indirect category lead; quarterly review	Head of contracts; monthly review	AP controller; monthly review

Each mandate is short enough to read in a minute and specific enough to audit. Set thresholds jointly with finance, legal and internal audit, and publish them alongside the autonomy register described in Chapter 6.

WHEN MANDATES COLLIDE

Which mandate prevails when two of them conflict?

Exhibit 3.4 The lowest bid comes from a supplier the risk agent is concerned about



Conflicts between mandates are not defects; they are strategy trade-offs made visible. Set the priority order in advance and in writing. Expressed this way, a category strategy stops being an annual slide deck and starts running every hour.

CHAPTER 04

Inside the Machine

The reliability of an always-on agent is an infrastructure problem before it is a model problem. The reference architecture below shows the five layers every deployment needs, and the two capabilities that span all of them.

Exhibit 4.1 Reference architecture of an always-on S2P agent

Signals	ERP and P2P events · intake · contract dates · indices · FX · ratings · sanctions · news
1 Sense	Event, schedule, threshold and anomaly triggers
2 Model	Knowledge graph linking suppliers, contracts, spend and obligations
3 Reason	Tiered cognition, policy checks and a decision core
4 Act	Scoped tools, SAP and Oracle write-back, supplier messages, undo
5 People	Daily digest, escalations, approvals above authority, mandate changes

Memory spans every layer

Governance mandates, audit log, stop controls

SENSING

Four kinds of trigger cover almost everything that matters in source-to-pay

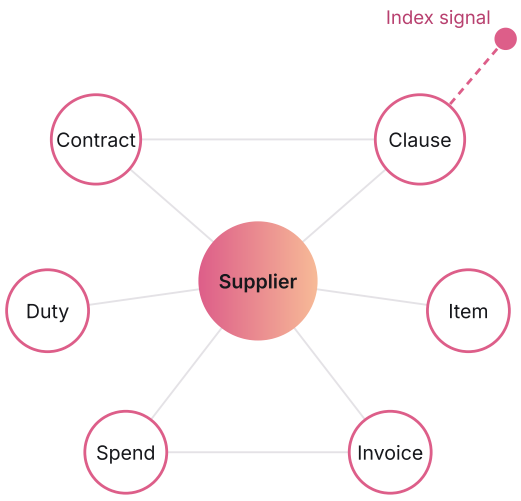
Exhibit 4.2 Trigger types and examples

TRIGGER	MEANING	EXAMPLES
Event	Something happened	Invoice posted, goods received, intake request filed, bank details changed
Schedule	Time has passed	Notice window opens, rebate period closes, business review falls due
Threshold	A limit was crossed	Index moves outside the repricing band; spend crosses a rebate tier
Anomaly	Something looks wrong	Possible duplicate invoice; unusual unit price; delays clustering at one supplier

THE WORLD MODEL

When a price index falls, the agent must know immediately which contracts carry an index clause, which items they cover and which suppliers are affected. Relational tables owned by separate modules cannot answer that quickly. A procurement knowledge graph can, because one signal reaches everything it touches in a single step.

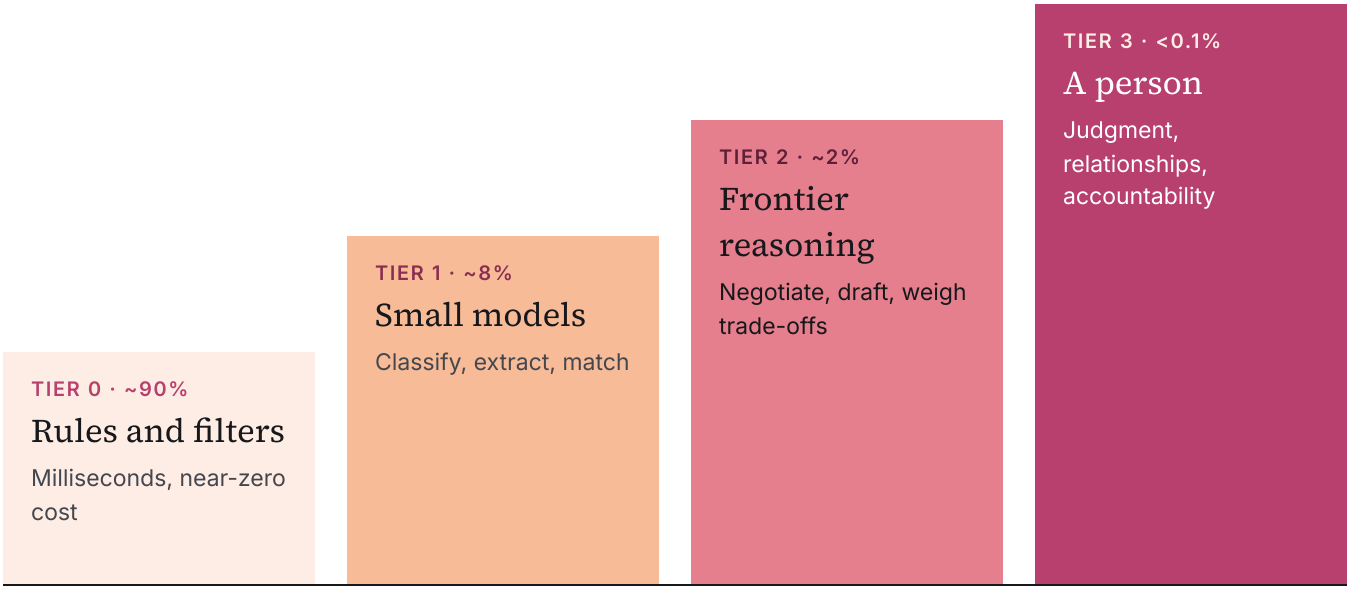
Exhibit 4.3 One signal, one step



TIERED COGNITION

Escalate reasoning the way you escalate people

Exhibit 4.4 The escalation staircase (shares of decisions are illustrative)



MEMORY THAT COMPOUNDS

EPISODIC

What happened

The last three negotiations with this supplier, and what moved them.

SEMANTIC

What is true

Policies, clause meanings, approved suppliers and tolerances.

PROCEDURAL

How we work

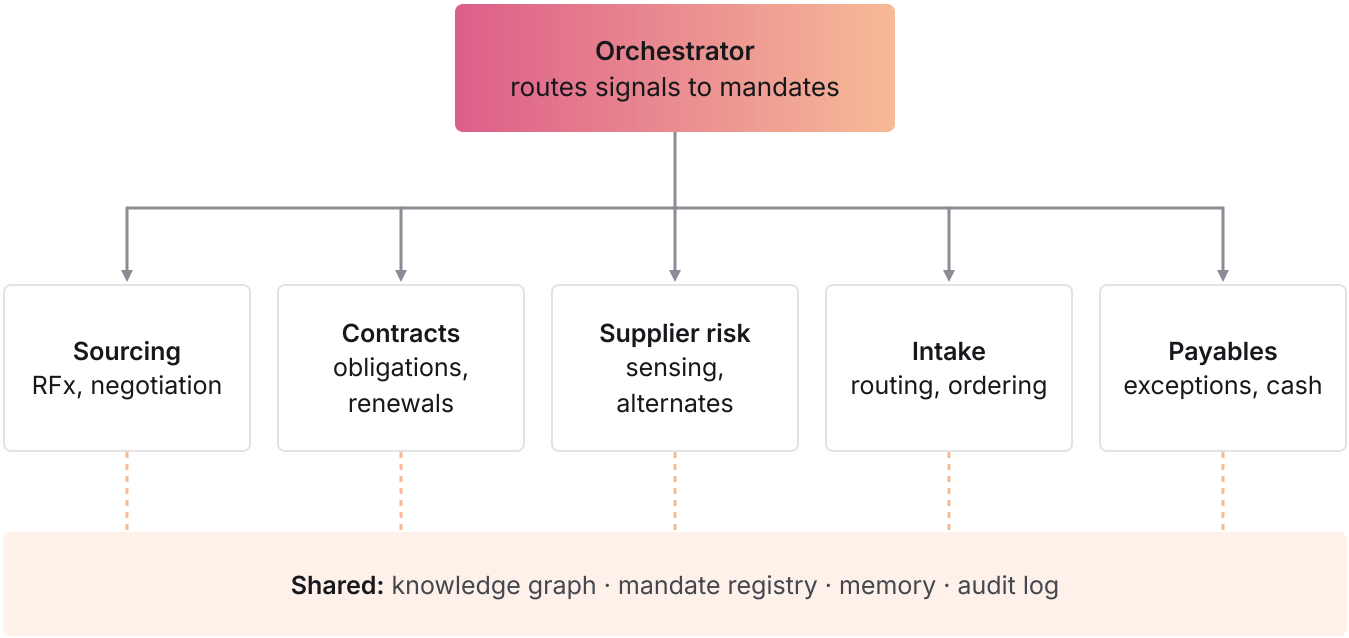
The negotiation playbook and escalation norms, refined by outcomes.

Design forgetting deliberately. Expired contracts, retired policies and one-off exceptions should age out; otherwise the agent applies yesterday's rules with confidence.

AGENT TOPOLOGY

One orchestrator, specialist agents and a shared world model

Exhibit 4.5 Multi-agent topology



THE ACTION LAYER: FOUR NON-NEGOTIABLES

01

Scoped permissions

Each agent receives the narrowest set of tools its mandate requires.

02

Idempotent write-back

Posting to the ERP twice must have the same effect as posting once.

03

Compensating actions

Every action has a defined undo: recall, reverse or withdraw.

04

Segregation of duties

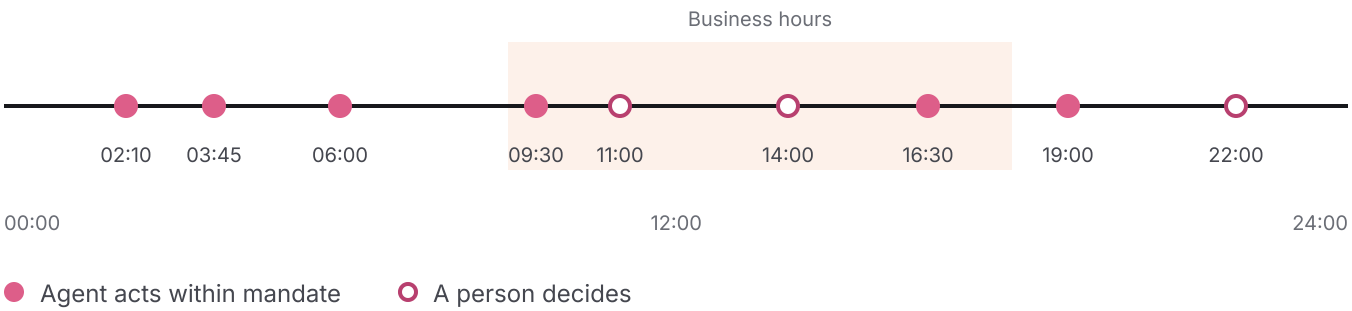
The agent that raises a commitment cannot approve or pay it.

CHAPTER 05

Twenty-Four Hours in an Autonomous Procurement Function

Meridian Components is a fictional $8 billion industrial manufacturer with 6,000 suppliers. Follow one ordinary Tuesday and notice how much happens outside office hours.

Exhibit 5.1 Nine moments in one day



Four of the nine moments occur before the first buyer logs in. The next two pages describe each one: what the agent did, what a person did, and which measure moved. The company, figures and events are illustrative.

02:10 – 09:30

Before the office opens, four matters have already been handled

02:10 A resin index falls 7% overnight

AGENT	PERSON	MEASURE
Identifies 14 contracts with index clauses and drafts repricing requests	None required	Price-adjustment capture

03:45 A tier-two supplier’s credit is downgraded

AGENT	PERSON	MEASURE
Maps exposure across 31 parts and pre-qualifies two alternates	Category lead reviews the briefing at 08:15	Time to mitigation

06:00 The overnight invoice batch arrives

AGENT	PERSON	MEASURE
Clears price variances against the amended contract and requests one credit note	Two genuine exceptions are routed to payables	Touchless rate; exception age

09:30 An engineer requests 40 torque sensors

AGENT	PERSON	MEASURE
Finds an existing contract and catalog item, then places the order	Engineer confirms quantity and site	Spend under contract; cycle time

11:00 – 22:00

People spend the day on decisions only people can make

11:00	**Packaging tender closes above authority** The agent secured a price 6% below benchmark; the $72k award exceeds its limit, so the category lead approves it.	PERSON
14:00	**Logistics contract 90 days from renewal** The agent issues a non-renewal notice and compiles performance data; the manager leads the renegotiation.	PERSON
16:30	**A rebate tier is crossed** Spend passes the contracted volume tier; the agent calculates the rebate and files the claim.	AGENT
19:00	**Cash position allows early payment** The agent accelerates 38 approved invoices that offer 2/10 terms, within treasury policy.	AGENT

Exhibit 5.2 22:00 digest to the CPO (illustrative)

Autonomous actions	212
Value captured or protected	$186k
Decisions referred to people	11
Agent actions overridden	1
Pending for tomorrow	3

Eleven decisions made by people all day. Each of them mattered.

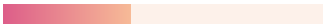
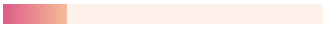
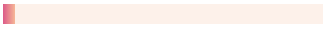
Some capacity is released. In most functions it is absorbed by rising workload, which Hackett expects to grow 8% in 2026, and redeployed into the supervisory roles described in Chapter 8. The rest is spent where judgment, relationships and accountability are required.

CHAPTER 06

Earned Autonomy

Governance for agents that never clock out. Autonomy is granted per decision class, earned through track record and withdrawn as soon as performance drifts.

Exhibit 6.1 The autonomy ladder

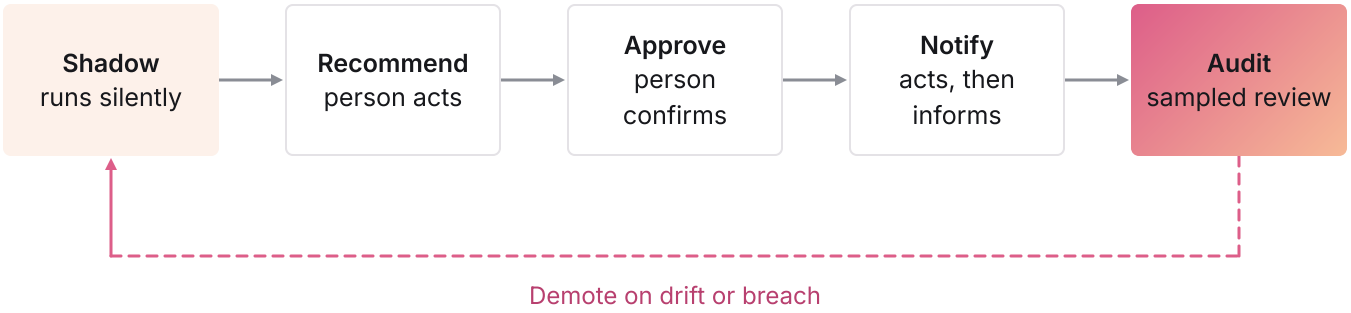
L5	Autonomous in mandate	Acts; reported in aggregate	
L4	Act and audit	Acts; actions sampled afterward	
L3	Act and notify	Acts; a person can reverse	
L2	Act with approval	Prepares; a person approves	
L1	Recommend	Proposes; a person acts	
L0	Observe	Watches and reports only	

Levels belong to decisions, not to agents. The same agent can operate at L5 when clearing a $40 freight variance and at L1 when a supplier termination is on the table.

THE PROMOTION PATH

Agents earn their way up the ladder, one decision class at a time

Exhibit 6.2 From shadow mode to audited autonomy



Promote when accuracy holds for an agreed number of cycles, overrides stay below tolerance and no control has been breached.

THE AUTONOMY REGISTER

This is the governance record the CFO and auditors will ask for. If a decision class is not in the register, the agent may not act on it.

Exhibit 6.3 Sample register entries (illustrative)

DECISION CLASS	LEVEL	TOLERANCE	OWNER	REVERSED
Clear price variance on PO invoice	L5	≤ 2% and ≤ $500	AP controller	0.4%
Award tail-spend tender	L4	≤ $50k, three bids	Indirect lead	1.8%
Issue non-renewal notice	L3	Non-critical suppliers	Head of contracts	3.1%
Switch to alternate supplier	L2	Pre-qualified only	Supply risk lead	12% rejected
Terminate a supplier	L1	Never alone	CPO	—

FAILURE MODES

An agent that never sleeps can also be wrong around the clock. Design the controls first.

Exhibit 6.4 Failure modes specific to always-on agents

FAILURE	WHAT HAPPENS	CONTROL
Runaway loops	Two agents keep triggering each other's requests and counters	Rate limits and loop detection per counterparty
Cascading actions	One incorrect master-data change spreads into hundreds of orders	Blast-radius limits and staged rollout
Silent drift	Market conditions shift and yesterday's tolerances stop making sense	Drift monitors that trigger demotion
Poisoned documents	An invoice or proposal carries hidden text instructing the agent	Documents treated as data, never as instructions
Adversarial agents	A supplier's agent probes for your walk-away price	Varied tactics and sealed reservation values

CIRCUIT BREAKERS REQUIRED ON DAY ONE

Global stop

One action pauses every agent and preserves state.

Mandate pause

Freeze one mandate without affecting others.

Value ceiling

A daily cap on autonomous commitments per agent.

Anomaly trip

Unusual volume demotes the agent to L2 automatically.

Gartner expects more than 40% of agentic AI projects to be canceled by the end of 2027, with inadequate risk controls among the leading causes.

CHAPTER 07

The Economics of Never Sleeping

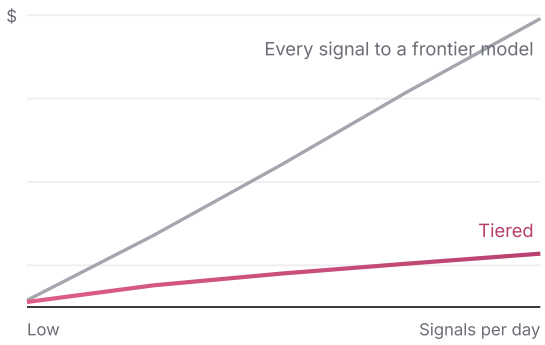
The cost structure moves from labor to compute and attention. Leaders manage what it costs to wake an agent, and what it costs to wake a person.

Exhibit 7.1 The value equation

Value	=	Signals	×	Act rate	×	$ / action	−	Compute	−	Interrupts
NET		COVERAGE		AUTONOMY		MANDATE QUALITY		TIERING		HUMAN TIME

Continuous sensing is inexpensive; continuous reasoning is not. An architecture that sends every signal to a frontier model will see cost scale with the noise in the data rather than the value in it. Tiered cognition keeps the curve flat.

Exhibit 7.2 Daily compute cost against signal volume (illustrative)



THE PRICING SHIFT

PAST

Per seat

Pays for logins, which makes little sense when most work is done by software nobody logs into.

TRANSITION

Consumption

Pays per action or signal; fair, but it penalizes noisy data unless tiering is built in.

NEXT

Outcome-linked

A base fee plus a share of verified value, aligning the vendor with your results.

A DEFENSIBLE BUSINESS CASE

Year-one value for an enterprise with $2 billion of addressable spend

Exhibit 7.3 Year-one value bridge, base case, $ million (illustrative)

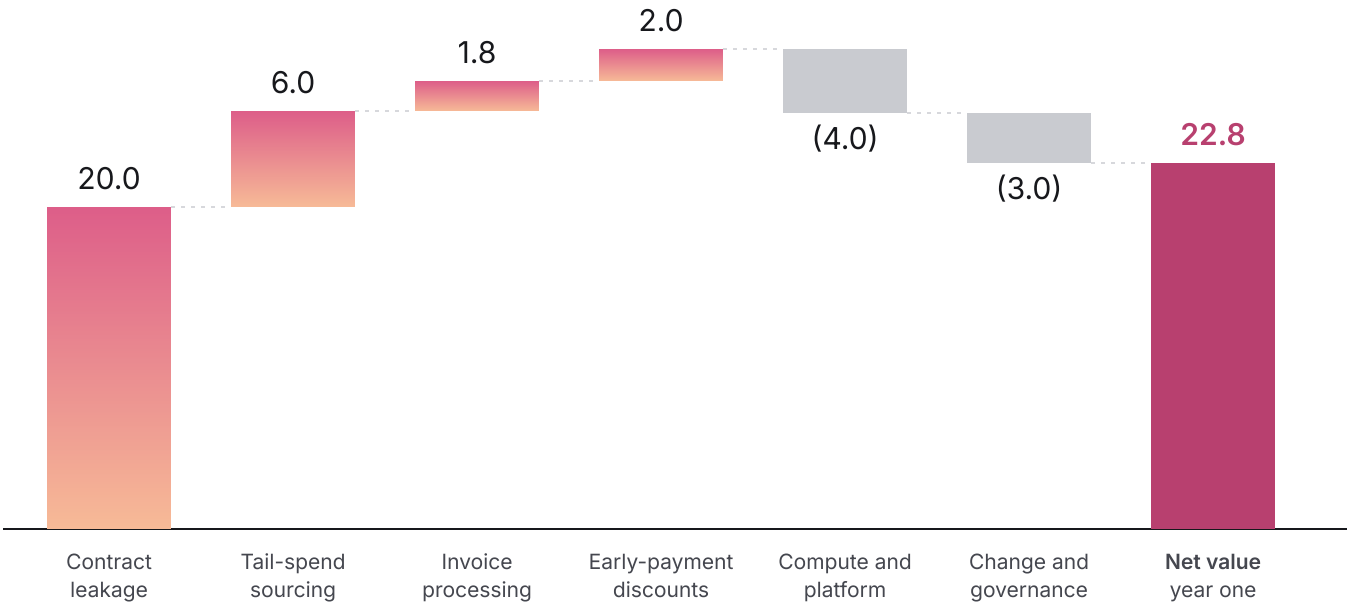


Exhibit 7.4 Assumptions behind each line, $ million

LINE	BASIS	LOW	BASE
Contract leakage	0.5–1.0% of addressable spend recovered; WorldCC models 2–3% in year one	10.0	20.0
Tail-spend sourcing	Tail is 20% of spend; half newly competed at 1.5–3% savings	3.0	6.0
Invoice processing	500,000 invoices; 25–50% of the $7.23 gap to best-in-class closed in year one	0.9	1.8
Early-payment discounts	2% captured on an additional $50–100 million of eligible spend	1.0	2.0
Compute and platform	Vendor fees and inference cost	(5.0)	(4.0)
Change and governance	Program team, audit and training	(3.0)	(3.0)
Net year-one value		**6.9**	**22.8**

CHAPTER 08

The Supervisory Procurement Organization

Procurement teams stop processing work and start supervising autonomous capacity. The scarce skill becomes designing mandates, not executing transactions. Capacity released by agents is largely absorbed by rising workload; the rest moves into four new roles.

Exhibit 8.1 Four new roles

Mandate architect

Translates category strategy into testable, auditable mandates.

Agent supervisor

Runs the exception queue and owns the interrupt budget.

Data steward

Keeps suppliers, clauses, items and hierarchies trustworthy.

Autonomy auditor

Samples actions, maintains the register and approves promotions.

Exhibit 8.2 How existing roles change

ROLE	TODAY	WITH ALWAYS-ON AGENTS
Category manager	Runs sourcing events	Writes and tunes mandates
Buyer	Raises purchase orders	Handles the complex minority of requests
Payables specialist	Keys and matches invoices	Resolves genuine disputes
CPO	Reviews reports	Governs autonomy

NEW MEASURES

Cycle time and cost per invoice still matter. These five show whether autonomy is working.

Exhibit 8.3 The autonomy scorecard

MEASURE	DEFINITION	CALCULATION	SHARPENS
Signal-to-action latency ↓	Time from a signal appearing to the first action, by agent or person	Median of action time minus signal time	Cycle time
Autonomous value share ↑	Share of captured value delivered by agents within mandate	Agent value ÷ total value	Savings realized
Interrupt rate ↓	Referrals to people per thousand signals processed	Referrals ÷ signals × 1,000	Touchless rate
Override rate ↓	Share of agent actions reversed by a person; a leading indicator of drift	Overrides ÷ autonomous actions	Exception rate
Mandate coverage ↑	Share of addressable spend governed by an active, reviewed mandate	Mandated spend ÷ addressable spend	Spend under management

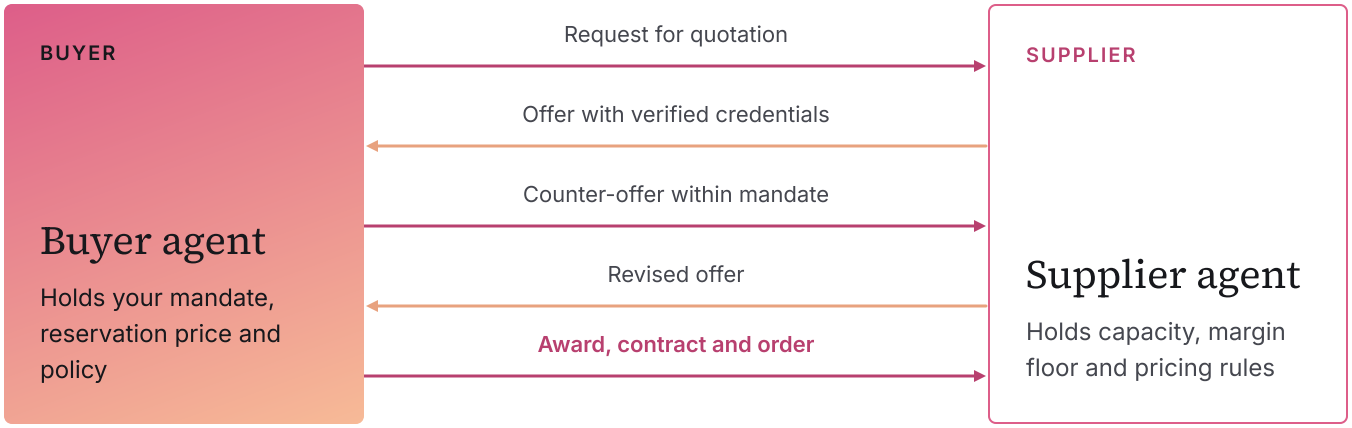
Teams that re-check every agent action never recover the capacity. Publish the override rate weekly, recognize the catches, and let the evidence, rather than anxiety, decide when to promote.

CHAPTER 09

The Road to 2030

Gartner forecasts that 90% of B2B buying, more than $15 trillion of spend, will be intermediated by AI agents by 2028. Once suppliers run agents too, procurement becomes a market that clears at machine speed.

Exhibit 9.1 A negotiation between agents



NEW STRATEGIC RISKS

Speed asymmetry

A supplier agent that reprices hourly will outpace a buyer that reviews quarterly.

Pattern leakage

Predictable concessions teach the other side where your walk-away point lies.

Trust infrastructure

Without verifiable credentials, a genuine supplier agent cannot be told apart from an impostor.

BEFORE YOU SCALE

Six hard problems every leadership team will face

The frameworks in this book are necessary but not sufficient. These six problems decide whether a promising pilot becomes an operating model. The first three are covered here; the remaining three follow.

1 Legal authority and disclosure

THE PROBLEM

When an agent issues a non-renewal notice or accepts an offer, someone has made a commitment. Delegated authority, contract law and, in the EU, the AI Act's transparency obligations, which apply from August 2026, all bear on who that someone is and what the counterparty must be told.

WHAT TO DO

Map every decision class in the register to a formal delegation of authority. Disclose agent involvement in supplier communications, and have counsel approve every template an agent may send.

2 The data you actually have

THE PROBLEM

The knowledge graph in Chapter 4 can sound like a multi-year master-data program. Waiting for perfect data is the most common reason agent programs stall.

WHAT TO DO

Build the minimum graph for the first three mandates only: the suppliers, contracts and spend those mandates touch. Expand mandate by mandate, and let agent errors show where data quality matters most.

3 Measuring the Attention Gap

THE PROBLEM

The 90-day plan assumes a baseline that most functions do not have. Without one, neither the business case nor the results can be defended.

WHAT TO DO

Sample the last 90 days of renewals, index movements and invoices with discount terms. For each, record when the signal appeared, when someone acted and what value was lost. That ledger becomes both the baseline and the business case.

BEFORE YOU SCALE, CONTINUED

Direct spend, suppliers and model risk

4 Direct versus indirect spend

THE PROBLEM

Most early examples, including many in this book, come from indirect and tail spend. Direct materials involve production schedules, qualified suppliers and multi-year contracts, where a wrong action stops a line.

WHAT TO DO

Start in indirect spend, where volume is high and risk is low. In direct spend, deploy agents first for sensing and preparation at L1 or L2, with planners and category leads making the decisions.

5 Supplier reaction

THE PROBLEM

Strategic suppliers may resent negotiating with software, and relationships built over years can erode quickly if automation feels adversarial.

WHAT TO DO

Begin with tail suppliers, disclose agent involvement and always offer a named human contact. Track supplier satisfaction alongside savings, and review it with the same rigor.

6 Model risk before promotion

THE PROBLEM

The circuit breakers in Chapter 6 catch failures in production. They do not show that an agent is fit to be promoted in the first place.

WHAT TO DO

Test each decision class against historical cases with known outcomes, red-team it with adversarial documents, and re-validate whenever the underlying model, prompts or policies change.

FOUR PREDICTIONS

What I expect to be true by the end of the decade

Exhibit 9.2 The author's forecasts Confidence

2027	**The agent-washing shakeout** As canceled projects mount, buyers judge products by the actions they take rather than the label they carry. Vendors that only recommend lose ground.	■■■■■
2028	**Autonomy becomes a separately priced tier** Suites charge for agents that act, and outcome-linked fees appear in a meaningful share of enterprise procurement contracts.	■■■■□
2029	**The autonomy register is audited** External auditors request it alongside financial controls, and mandates become a standard governance object.	■■■□□
2030	**Agent-to-agent is the default for routine spend** Routine categories clear between buyer and supplier agents, verified credentials gate supplier onboarding, and teams are sized by the decisions they govern.	■■■□□

The timeline is debatable; the direction is not. Each year of delay is a year of mandates that competitors have refined and you have not.

THE FIRST YEAR

Three phases, each with a clear exit test

Exhibit 9.3 A 90-, 180- and 365-day plan

90

DAYS · FOUNDATIONS

See the gap

Measure today's losses and connect the signals.

- Size the Attention Gap with your own data
- Stream ERP, P2P and contract events
- Select three low-risk, high-value mandates
- Run agents in shadow mode only

EXIT TEST

Shadow agents would have caught value that was actually lost.

180

DAYS · FIRST AUTONOMY

Earn trust

Let agents act on narrow decision classes, visibly.

- Promote two decision classes to L3
- Publish the register and circuit breakers
- Start the daily digest and override review
- Appoint a mandate architect and a supervisor

EXIT TEST

Override rate within tolerance for eight consecutive weeks.

365

DAYS · SCALE

Run always-on

Extend mandates across S2P and change how success is measured.

- Cover 40% of addressable spend with mandates
- Add autonomy measures to the CPO scorecard
- Move vendors toward outcome-linked pricing
- Pilot agent-to-agent negotiation with two suppliers

EXIT TEST

Autonomous value share is reported to the board.

APPENDIX A

Twenty questions before the first mandate

Score each statement 0 (not true), 1 (partly true) or 2 (fully true). The maximum score is 40.

Data

01 Supplier master data has one owner and one identifier per supplier

02 Contracts are digitized, with key dates and clauses extracted

03 Spend is classified to a taxonomy at line level

04 ERP and P2P events can be streamed, not only exported in batches

05 External signals such as indices and ratings are licensed and accessible

Process

06 Approval thresholds are documented and applied consistently

07 Category strategies state explicit objectives and limits

08 Exception types are cataloged, each with an owner

09 Current touchless and discount-capture rates are known

10 Renewal and notice dates are tracked for all material contracts

Governance

11 Internal audit has reviewed an AI use case end to end

12 Segregation-of-duties rules can be applied to software agents

13 A named executive owns agent autonomy

14 Every system action can be logged with who, what and why

15 Legal has a position on commitments made by agents

People

16 Leaders present AI as added capacity, not headcount reduction

17 At least one person can write a testable mandate today

18 IT and procurement share one AI roadmap

19 Teams are measured on outcomes rather than activity

20 Key suppliers know that agents may contact them

0–16

Fix data and thresholds first; run in shadow mode only.

17–30

Ready for two or three narrow mandates at L2 or L3.

31–40

Scale across S2P and begin the 365-day plan.

APPENDIX B

Ten questions for vendors

01 Which events and external signals do your agents subscribe to as standard?

02 What is the median latency from an ERP event to agent evaluation?

03 Name five actions your agents take in production without a human click.

04 How are ERP write-backs made idempotent and reversible?

05 How are sensing and reasoning costs separated on the invoice?

06 Show the autonomy register and how a decision class is promoted.

07 Demonstrate the global stop and a single-mandate pause.

08 How do you defend against instructions hidden in supplier documents?

09 Share override rates from three production customers.

10 Will part of your fee be linked to verified value?

APPENDIX C

Glossary

Always-on agent
Software that senses continuously, holds state and acts within a mandate without being prompted.

Attention Gap
Value lost between a signal appearing and a person acting on it.

Mandate
Written standing intent: objective, scope, authority, constraints, escalation and review.

Earned autonomy
Granting action rights per decision class based on measured track record.

Autonomy register
The audited list of decision classes, levels, owners and tolerances.

Interrupt economics
Signal absorbed for each unit of human attention consumed.

Tiered cognition
Escalating from rules to small models to frontier reasoning to people, as needed.

Compensating action
The defined undo for an agent action.

BACK MATTER

Sources and notes

1 Gartner press release: over 40% of agentic AI projects will be canceled by end of 2027; about 130 of thousands of agentic AI vendors are genuine; 15% of day-to-day work decisions autonomous by 2028. June 25, 2025.
gartner.com/en/newsroom/press-releases/2025-06-25-gartner-predicts-over-40-percent-of-agentic-ai-projects-will-be-canceled-by-end-of-2027

2 Gartner press release: 40% of enterprise apps will feature task-specific AI agents by 2026, up from less than 5% in 2025. August 26, 2025.
gartner.com/en/newsroom/press-releases/2025-08-26-gartner-predicts-40-percent-of-enterprise-apps-will-feature-task-specific-ai-agents-by-2026

3 Gartner forecast that AI agents will intermediate 90% of B2B buying, and more than $15 trillion of purchases, by 2028. Reported by Digital Commerce 360, November 2025.
digitalcommerce360.com/2025/11/28/gartner-ai-agents-15-trillion-in-b2b-purchases-by-2028

4 The Hackett Group, 2026 Procurement Key Issues: workload up 8% while headcount and budgets decline. March 2026.
thehackettgroup.com/the-hackett-group-reports-rapid-progress-in-procurements-ai-agenda

5 World Commerce & Contracting: 8.6% of contract value lost on average to poor contract management. August 2025.
worldcc.com/resource/contract-management-an-overlooked-driver-of-business-agility-and-financial-performance

6 WorldCC and Ironclad: 11% of contract value lost post-award; missed price adjustments 1–2% of spend; 2–3% recoverable in year one. Reported by Digital Journal.
digitaljournal.com/business/contracts-signed-value-lost-how-businesses-are-leaking-11-of-spend

7 Ardent Partners 2026: $9.90 average vs $2.67 best-in-class cost per invoice; 19.9% vs 11.8% exception rate. As summarized by Tailride, October 2026.
tailride.so/blog/invoice-management-statistics-2026

8 Regulation (EU) 2024/1689, the Artificial Intelligence Act, Article 50 on transparency obligations.
eur-lex.europa.eu/eli/reg/2024/1689/oj

The 37% figure is a calculation: forgoing a 2% discount to pay 20 days later equals a return of 2 ÷ 98 × 365 ÷ 20. Benchmarks vary by survey year; each figure is dated to its source.

Exhibits 1.5, 2.2, 3.3, 4.4, 5.1, 5.2, 6.3, 7.2, 7.3 and 7.4, and the company described in Chapter 5, are illustrative. Exhibit 9.2 and the lessons from consumer agents in Chapter 2 present the author's own views.

Markets move overnight. Procurement should be able to respond overnight too.

FOUR IDEAS FROM THIS BOOK

The Attention Gap	Where value is lost between a signal and an action
Mandates, not prompts	Standing intent that agents can hold and auditors can test
Earned autonomy	Action rights granted per decision, on the evidence of track record
Interrupt economics	Signal absorbed for every minute of human attention