

From Intelligent ERP to Autonomous Supply Networks

The Future of Supply Chain Management in the Age of AI

The strategic question is not how many tasks AI can perform.
It is which decisions an organization can safely delegate.

Prepared by Ahmed Abdulrahman Alobaid

أحمد عبدالرحمن العبيد

procurement specialist

Revised research edition

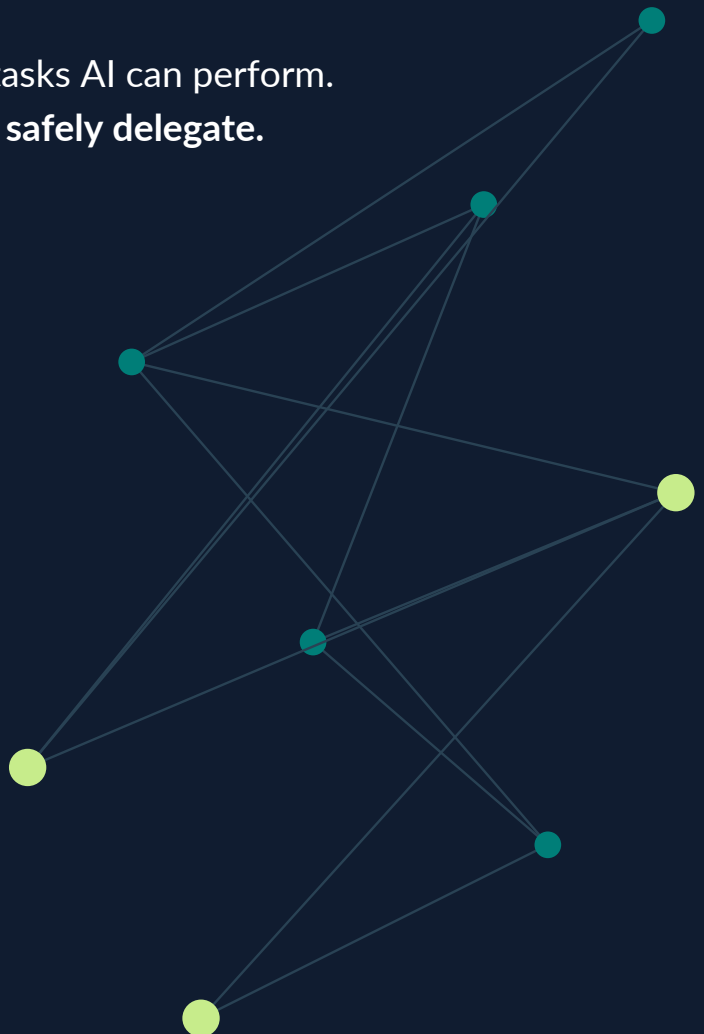
ORACLE FUSION BENCHMARK

GLOBAL PLATFORMS • SAUDI ARABIA & GCC

Verified capabilities, control architecture,
conditional scenarios and strategic implications.

18 SEPTEMBER 2026

Research report • 51 sources



ANALYSIS

Executive summary

ERP will remain central. The larger change is who makes routine decisions and who answers for them.

Between 2026 and 2035, the strongest case is for delegating selected decisions under explicit limits. Agents would gather evidence and coordinate work; forecasting models and optimizers would calculate the options; enterprise systems would record the resulting commitments. ERP therefore remains the authoritative record. This report expects a decision-coordination layer to develop around it, while planning, procurement and financial controls continue to matter.

Parts of this model already appear in product documentation. Oracle describes autonomous sourcing for eligible purchases. Microsoft's supplier-communication feature, however, is still identified in the reviewed documentation as a production-ready preview. The distinction matters when choosing a pilot: these products offer specific capabilities, with different limits, rather than unrestricted network-wide execution. [S07, S13]

Where leadership decisions matter

Set authority before changing the interface. Decide what an agent can read, recommend, communicate and commit. Saving a buyer's preparation time will have limited effect on turnaround if every request still waits in the same approval queue.

Follow the effect through operations and finance. Measure compliant completion, exception work, landed cost, service and cash. Lower prices can create shortages, and longer payment terms can weaken suppliers. Keep proposed savings, working-capital release and annual profit separate when assessing the result.

Choose where delegation makes sense. Repetitive work with observable outcomes and a recovery path is an earlier candidate. Critical sourcing, scarce-supply allocation, product substitutions, sensitive supplier relationships and uncertain trade classifications warrant more judgment. Expansion depends on usable data, prepared counterparties and recovery arrangements as well as model quality.

INITIAL SCOPE

Begin with one procurement workflow and one planning exception. Extend execution rights after shadow tests, restricted live use and independent review of the outcomes.

For Saudi Arabia and the GCC, the near-term opportunity lies in coordinating enterprise decisions with existing trade and conformity services. ZATCA documents AI-supported validation in a particular Fasah declaration workflow. That specific service should not be read as evidence that customs clearance has become autonomous. [S28]

RELEASE DETAIL THAT CHANGES THE DECISION

Oracle's 26C inventory command center excludes lead-time and upstream supply constraints from recommendations. A separate feasibility check is needed before committing the action. [S41]

READING GUIDE

Contents

Management path: summary → toolkit → scenario → roadmap → action plan.

Executive summary	2
Evidence framework and methodology	4
Five-stage evolution and what comes after ERP	5
Oracle current-state benchmark and sourcing loop	7
Competitive landscape	11
What Supply Chain Professionals Can Use Today – 2026 Toolkit	12
Procurement and strategic sourcing workflows	13
Planning and inventory workflows	14
Logistics and import/export workflows	15
General-purpose AI workbench and starter briefs	16
Procurement, PPV and agent negotiation	17
Planning, inventory and six-agent disruption scenario	20
Conversational analytics and agent KPIs	23
Logistics, HS and import/export	25
Saudi Arabia and GCC operating models	26
Data, cybersecurity and accountability	28
People, economics and counterarguments	31
2026–2035 roadmap and scenarios	34
Organizational and professional action plans	36
Conclusions and leadership decisions	38
Function transformation matrix	39
Acceptance testing protocol	41
Decision engines, regional integration and Oracle release limits	42
Worked visual case: PPV versus enterprise value	45
2026 Emerging Supply Chain AI Research	46
References and evidence limitations	47

QUICK PRACTICAL ROUTES

Pages 12–16: current tools and workflows. Page 22: six-agent food/QSR disruption. Page 44: Oracle 26C limitations. Page 45: why a lower price can destroy value.

EVIDENCE FRAMEWORK

How to read the evidence

The evidence labels separate published functions, provider claims, research findings and this report's forecasts.

The evidence cutoff is 18 September 2026. This synthesis does not certify products, offer a legal opinion or estimate causal business effects. The earlier research edition had 43 references; the intervening toolkit edition had a different set of 48. Restoring three omitted studies brings this register to 51. The reconciliation below explains the changes.

Evidence label	What it establishes	What it does not establish
DOCUMENTED	A release or service describes a defined function.	Enabled in every tenant, implementation quality or independent ROI.
VENDOR-REPORTED	A provider describes availability, capability or a deployment result.	An audited comparison or representative causal effect.
EMERGING	Preview functions, engineering methods or research findings.	General availability or production fitness across use cases.
FORECAST / ANALYSIS	A reasoned scenario, operating design or recommendation.	A promised date, adoption rate or empirical observation.
ILLUSTRATIVE	Explicit synthetic assumptions used to test a decision or calculation.	A company's actual data, industry average or expected return.

Verification and boundaries

Oracle's release notes were used to check selected entries in its feature index. Microsoft's reviewed supplier-communication setup is labeled preview; future Oracle 26D entries are excluded from the current benchmark. The comparison concerns functions and authority. Vendors were not tested against a common live workload. [\[S06–S09, S13–S14\]](#)

Government pages are cited for the services they describe. The 2021–2025 NIDLP plan supplies historical context. Customer-support studies and agent benchmarks inform the discussion, but cannot establish procurement job losses or live supply-chain error rates. [\[S22–S24, S28, S32\]](#)

THE MATERIAL EVIDENCE GAP

The main gap is long-run, independent evidence connecting autonomous decisions to reliable execution, full operating cost and realized business value. Product descriptions alone cannot fill it.

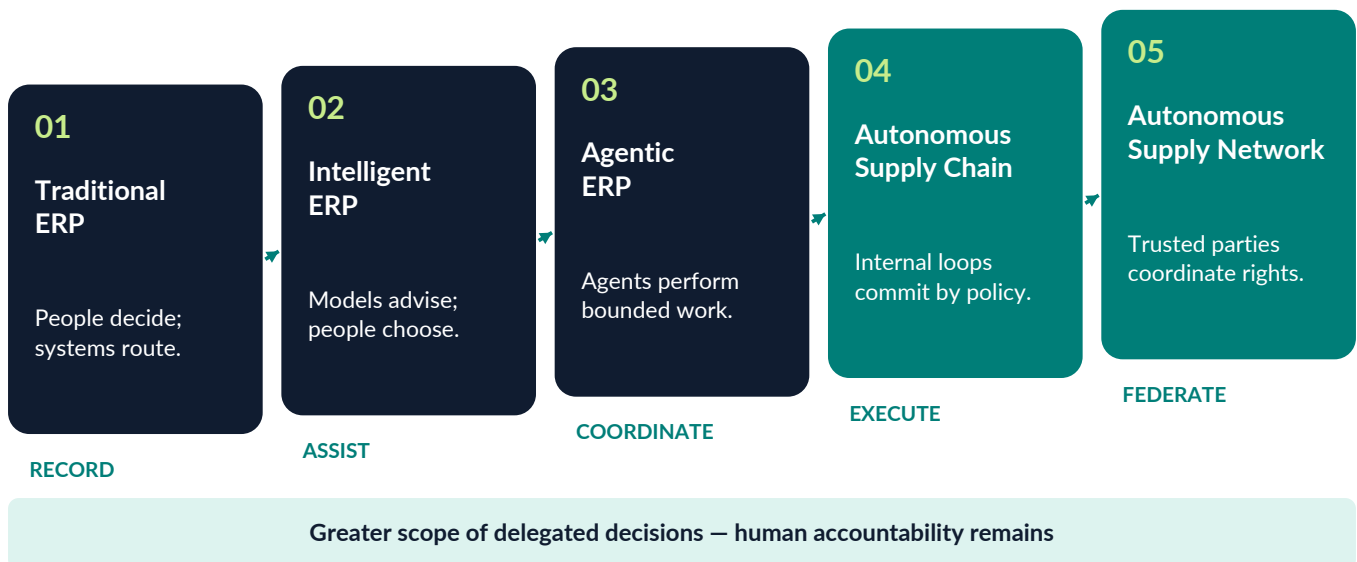
Reference reconciliation: 37 earlier references + four later Oracle release sources + seven toolkit sources + three restored studies = 51. Thus 40 of the earlier 43 are retained; three contextual items are not carried forward. Reference_Reconciliation.csv maps every earlier entry to its current ID or omission status.

QC review scope: S49–S51 were rechecked; other sources retain recorded provenance. Earlier S26/S32 access limits, the toolkit revision's unavailable S40, and the unverified SABER–Fasah link remain qualified in the register.

ANALYSIS

Five stages—not one technology leap

The stages distinguish the decisions delegated and the organizations involved.



The model begins with **traditional ERP**, where people interpret exceptions and systems record or route transactions. **Intelligent ERP** adds predictions and assistance. **Agentic ERP** can pursue an assigned objective across tools. Selected cross-functional decisions close without routine intervention in an **autonomous supply chain**; an **autonomous supply network** extends that coordination to independent firms. These definitions organize the analysis; they are not a vendor certification scheme.

A company need not pass through these stages in sequence. Its replenishment rules may already place routine orders automatically while an advanced assistant still requires approval for each action. Intelligence describes the system's capabilities; autonomy describes the discretion it has been given. One can increase without the other.

What changes between stages

To resolve a shortage, a system needs more than a convincing explanation. It must detect the problem reliably, work toward agreed objectives, identify a feasible response, obtain authority, execute and confirm the outcome. It also needs a recovery path. Across companies, this requires trusted identities, shared meanings and arrangements for resolving conflicting interests.

ASSESS THE WORKFLOW

Assess maturity by decision, product and relationship. Different workflows in the same business can reasonably sit at different stages.

Engineering guidance separates prescribed workflows from agents that choose their own steps. This distinction helps avoid treating routine automation and open-ended agency as the same design choice. [S18]

OPERATING-MODEL TRANSITION

What comes after the ERP interface

The next layer coordinates decisions; ERP continues to record and commit them.

01 Oracle Fusion today

Fusion records suppliers, items, stock, contracts and commitments. Selected agent functions add sourcing and assistance, subject to module and release availability. [\[S01–S09\]](#)

02 Agentic ERP

The agent gathers evidence, calls forecasting or optimization tools and works across modules. A separate policy service checks whether the proposed commitment is permitted.

03 Autonomous Supply Chain

Demand, inventory, procurement and logistics work from one versioned case. Routine eligible actions execute and reconcile; material exceptions go to an accountable reviewer.

04 Autonomous Supply Network

Buyer, supplier and carrier agents exchange verified offers and acknowledgments while retaining separate records, commercial limits and confidential objectives.

Oracle's current offerings already span intelligent ERP and selected agentic functions. The later stages extend coordination and authority; they do not imply that Oracle ends at stage 2 or that all firms follow one timetable.

The architecture that makes the transition operational

Execution starts with reliable records and events. Forecasting models estimate demand and uncertainty, and optimizers test the available choices. Agents bring those results together. Before any ERP, WMS, TMS, trade or finance transaction is committed, an independent policy service checks eligibility, budgets and segregation of duties. [\[S18–S21, S25\]](#)

Each case needs a traceable record: the input snapshot; model and policy versions; alternatives and constraint checks; approval tied to the amount, supplier and action; the tool call; and the execution receipt and reconciliation. Changed facts require renewed checks. Failed actions must be stopped or compensated. Agreement between two models does not replace this evidence.

WHAT CHANGES AFTER ERP

Work moves from separate screens to a shared decision case. Across firms, that case also needs verified counterparties, agreed terms, acknowledgments and a dispute process.

DOCUMENTED SCOPE + ANALYSIS

Oracle: the current transactional foundation

The benchmark spans Fusion Procurement, SCM and logistics—not one undifferentiated license.

Function	Established product scope	Boundary to test
Procurement / sourcing	Requisitions, purchasing, sourcing events, contracts, supplier qualification and collaboration. [S01]	Supplier/item eligibility, buying channels, delegated approvals and document quality.
Planning / S&OP	Demand and supply planning, scenario analysis and planning collaboration. [S02]	Constraint completeness, forecast error, solver settings, refresh frequency and adoption.
Inventory	Inventory visibility and execution, material movement and replenishment processes. [S03]	Unit of measure, location, ownership, lot status and physical-to-system accuracy.
Transportation	Transport planning and execution, carrier coordination and cost processes. [S04]	Coverage of counterparties, rates, service commitments and event freshness.
Global trade	Classification, screening, trade controls, landed-cost and customs-related preparation. [S05]	Jurisdiction, regulatory content, goods details, filing permissions and legal review.

Fusion already provides structured purchasing, planning and inventory workflows, including substantial rule-based automation. The practical contribution of generative AI is at the remaining gaps: interpreting unstructured requests, handling exceptions and carrying context between modules. Those are the areas an implementation should measure.

How to evaluate an actual Fusion environment

Begin with the actual tenant: its modules, quarterly release, regional availability, roles, integration accounts, supplier connections and retention settings. Record what is licensed, enabled and supported. This inventory often determines which of the publicly described features the team can sensibly test, and which require additional configuration or investment.

Trace a requisition until it becomes a binding purchase. Locate the budget check, the authoritative supplier record and the handoffs to planning and accounts payable. Resolve conflicting lead times or approved-supplier records before giving an agent execution rights. Otherwise, automation may transmit a known inconsistency more quickly to the next process.

IMPLEMENTATION CHOICE

Keep purchasing authority in Oracle where Oracle owns the transaction. Replicating controls in a separate agent application needs a specific, tested justification.

ANALYSIS

Oracle logistics and trade: a crucial boundary

Preparing a compliant action is different from controlling the external outcome.

Oracle Transportation Management covers transport planning and execution. Global Trade Management adds party screening, classification, license and permit controls, landed-cost analysis, trade-agreement support and customs-document preparation. Together they support important decisions and controls, while carriers, trading partners and authorities retain responsibility for their own actions. [S04, S05]

Three boundaries a future agent must respect

Commercial commitment. Track the difference between a rate quote, a reservation, a confirmed booking and an executed movement. The agent needs to know which state it is changing. After a timeout, it should check the existing booking before retrying and potentially creating another charge.

Operational reality. A port or warehouse may have no usable capacity despite a favorable arrival prediction. Late events, operating calendars and handling limits affect the choice. Replanning should show how old the evidence is and how uncertain the estimate remains, alongside the proposed schedule.

Regulatory authority. Preparing a classification suggestion or declaration does not confer the right to rule on it. Goods specifications, origin, valuation and the applicable jurisdictional measures still have to be established. Keep permission to prepare the document separate from permission to submit it through an authorized channel.

Example	Reasonable bounded authority	Escalation trigger
Late inbound shipment	Request evidence; compare approved alternatives.	Air-freight uplift exceeds budget or changes a customer commitment.
Document discrepancy	Reconcile values with authoritative records; prepare corrections.	Origin, value, classification or permit validity is disputed.
Carrier rate change	Compare contract and quote; seek confirmation.	New carrier, unusual liability terms or restricted route.

Internal planning becomes more useful when it can distinguish external estimates from accepted commitments. The operating view should make three things easy to see: what a counterparty has confirmed, what the system has inferred, and who can authorize the next action. A single status summary should not obscure those differences.

EXTERNAL HANDOFF

Use agents to shorten the evidence and communication chain. Retain the carrier’s, authority’s and approving manager’s separate decision rights.

DOCUMENTED + VENDOR-REPORTED

Oracle agents: inspect the release, not the slogan

Selected features documented by the cutoff; availability remains configuration-dependent.

Evidence / release	Examples reviewed	Interpretation
26A–26B release notes	Autonomous sourcing and award templates; approval-routing enhancements. [S07, S08]	A bounded sourcing workflow, not unrestricted commercial discretion.
Fusion AI release index	Planning exceptions, order release assistance, S&OP cycle assistance and inventory optimization entries through 26C. [S06]	Functions differ in release, data requirements and authority. The index is not a blanket activation list.
29 Jun 2026 announcement	Inventory Planning Command Center, Supplier Qualification Workspace, Production Readiness Workspace and Kanban Administrative Workspace. [S09]	Oracle announced four additional applications as available; operational results are vendor claims.
Not included as current	26D entries or unspecified later roadmap functions. [S06]	Future documentation is not counted as deployed capability.

What the evidence supports

The reviewed offerings extend beyond drafting text into workspaces and multi-step activities. Their names, however, reveal little about execution rights. A command center may prepare an action for a user to approve; a scheduled job may run unattended using prescribed logic. Assess the actual trigger, steps and approval points before assigning an autonomy label.

What an implementation team should request

For each feature, request its release, supported objects, trigger, allowed actions, failure handling, audit trail and approval points. Test denied actions, stale data and interrupted transactions. These reveal where the workflow needs human intervention, which a successful demonstration may not show.

Compare the feature with the workflow already configured in the business. Include work transferred to suppliers, exception reviewers and support staff when calculating time saved. Fewer clicks may improve usability, but a business case still needs evidence about total effort, decision quality and the cost of resolving failures.

EVIDENCE LIMIT

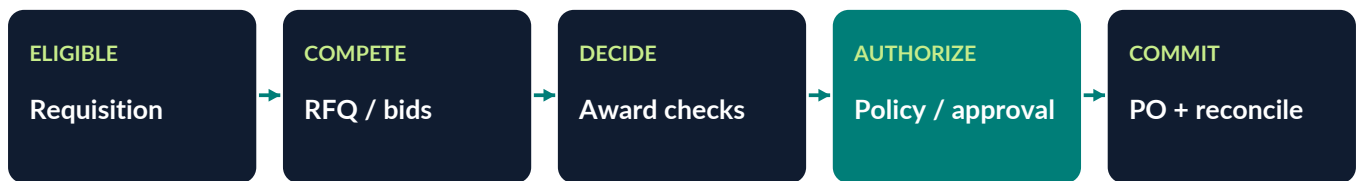
The release notes define what to test. The organization must still verify controls and outcomes in its own environment; no Oracle tenant was tested here.

ANALYSIS

A bounded Oracle sourcing loop

Current documentation illustrates why governance is part of the product—not an optional add-on.

Oracle's 26A templates use eligible requisition lines to create sourcing events, select suppliers from historical information and support awards and purchase documents. The documented focus is lower-dollar, high-volume indirect or tail spend. Configured policies and approvals apply, and processing can run through scheduled jobs. [S07]



Recommended control envelope: eligibility · evidence · exposure · identity · recovery

In 26B, autonomous-sourcing and autonomous-award attributes can be used in approval routing. A business can therefore delegate selected processing while retaining explicit review points for the awards or commitments it considers material. [S08]

Recommended operating design around that capability

Eligibility should cover the item and supplier, contract terms, currency, delivery destination, category and budget. Price alone is too narrow a gate. Keep regulated goods, new bank details, substitutions and concentrated critical supply outside the initial scope until their specific controls have been tested.

Bids need a common comparison basis: quantities, units, delivery terms and service obligations. A cheaper line can lead to a more expensive delivered purchase. Where evidence is missing, record the gap and route the exception; an invented non-price score would conceal the unresolved issue.

After award, reconcile the purchase document with supplier acknowledgments, receipts and invoices. Look for changed commitments, duplicate orders, canceled requests and mismatches. An unresolved disagreement should return to the buyer, rather than be settled by silently substituting the closest available alternative.

INITIAL DELEGATION

A sensible starting scope is RFQ preparation and routing, with award recommendations. Automatic commitment can be limited to repeat, preapproved purchases. The company sets the thresholds and formal mandate.

This shifts the buyer's effort toward category rules, constraints, fallback suppliers and exceptions. Commercial judgment is still present: it determines the policy used by the agent and resolves cases that the policy does not cover reliably.

PRODUCT EVIDENCE + ANALYSIS

Competitive landscape: different control points

These platforms overlap, but they do not occupy the same place in the architecture.

Platform	Center of gravity / evidence	Procurement decision implication
Oracle Fusion	Integrated transactional and SCM processes; selected released sourcing agents. [S01, S02, S03, S04, S05, S06, S07, S08, S09]	Exploit existing authority and records; verify module, release and control configuration.
SAP	Enterprise SCM suite; Joule planning assistance and agent-supported exception / inventory analysis. [S10, S11]	Useful where SAP owns the process; product description is not blanket GA verification.
Kinaxis Maestro	Planning orchestration combining AI, optimization, simulation and Maestro Agents. [S12]	Evaluate plan quality and response speed; keep commitments reconciled with ERP.
Microsoft Dynamics 365	Supplier-communication agent with separate identity; reviewed setup remains production-ready preview. [S13, S14]	An adjacent communications loop, not proof of autonomous strategic procurement.
project44	Visibility plus agentic logistics orchestration and execution-related workflows. [S15]	Test network coverage, booking authority and event reliability—not only tracking displays.
Pactum	Specialist policy-driven supplier negotiation. [S16, S17]	Evaluate category fit, counterparties, terms and realized value; connect awards to ERP controls.

Choose a platform against a defined decision: can it complete that workflow safely using the right records and existing authority? Agent count provides little help with this question. Additional agents may duplicate context and make coordination or incident diagnosis harder, so their contribution needs to be demonstrated.

A stable ERP may remain the best place to execute commitments, with a specialist planner handling constrained scenarios and a network platform supplying external events. Evaluate that combination as a whole. Integration work and fragmented controls can erase the advantage of an impressive standalone feature.

SELECTION METHOD

Give candidates the same normal, ambiguous, stale-data and execution-failure cases. Include unavailable suppliers and denied permissions. Compare compliant outcomes, intervention work and full operating cost.

This comparison identifies scope and architectural fit, not market share or performance rankings. Public release detail varies between vendors; procurement diligence must close the remaining gaps.

CURRENT TOOLS + RECOMMENDED WORKFLOWS

What Supply Chain Professionals Can Use Today

2026 Toolkit — practical starting points, not a shopping list.

An approved analysis workspace and a reliable operational system are enough to begin many of the workflows in this toolkit. Enterprise execution requires additional modules, data, permissions and configuration. The examples describe ways to organize that work; they should not be read as promises of ready-made integrations.

Team	Enterprise route to evaluate	Useful first outcome
Procurement	Oracle Fusion sourcing [S07–S08]; Dynamics 365 supplier communications (preview) [S13]	Compare quotes and chase missing acknowledgments
Planning	Oracle Planning; SAP Joule planning; Kinaxis Maestro [S02, S11–S12]	Explain demand exceptions and test feasible replenishment
Strategic Sourcing	Oracle sourcing / qualification; Pactum negotiation [S01, S16–S17]	Build a cited category dossier and risk-adjusted award comparison
Logistics	Oracle OTM Shipment Change Assistant [S38]; project44 [S15]	Prioritize late shipments and assemble approved recovery options
Import/Export	Oracle GTM [S05]; authorized Fasah / broker workflows [S28–S29]	Build a ready-to-review shipment dossier and discrepancy register

Choose the lightest implementation that closes the gap

Route	What to use it for	Entry condition
USE ON APPROVED FILES	ChatGPT, Claude, Copilot in Excel or eligible Gemini in Sheets for analysis and preparation. [S42–S47]	Approved data scope; a human reviews evidence, calculations and outputs.
CONFIGURE WITH IT	Native enterprise agents and workflow routing such as Power Automate approvals. [S07, S13, S48]	Tenant readiness, dedicated identity, correct reviewers and recorded execution states.
PILOT, THEN DELEGATE	A cross-tool agent that executes a narrow repeatable decision loop.	A signed mandate, representative tests, rollback and measured net value.

A USEFUL START THIS WEEK

Take one recurring decision and a small approved sample. Compare the accepted AI-assisted output with the current process, including the time spent checking and repairing it. That is the useful baseline.

PRACTICAL WORKFLOWS

Procurement and strategic sourcing

Turn approved evidence into a comparable bid, then a defensible award.

1. Quote comparison and purchase-order follow-up

ChatGPT or Claude can support analysis of approved quotations and documents; Copilot in Excel can help maintain the review workbook. For enterprise processing, assess Oracle sourcing or Dynamics 365 supplier communications. The latter remains a preview in the reviewed documentation. [S07, S13, S42, S45–S46]

Bring specifications, quotations, PO lines and supplier masters into one comparison. Preserve the source of each value while normalizing currency, case packs, units, tax treatment and delivery terms. Leave gaps and contradictions visible. The immediate outputs are a comparable bid sheet and draft confirmation requests for the buyer to review.

Before anything is sent or committed, the buyer resolves product equivalence, exclusions and unusual terms. Record the reviewer's response and the resulting PO state. Power Automate can route this sign-off, but the purchasing system must still enforce the authority to execute. [S48]

2. Category research, should-cost and negotiation preparation

For category work, combine cited commodity and supplier research with a validated should-cost workbook. Oracle sourcing or a specialist such as Pactum can support a separately authorized commercial process when that workflow fits the category. [S01, S16–S17, S43]

Connect spend to contract index clauses, published input-price series, capacity evidence and switching lead times. The resulting dossier should show exposures and a range of landed-cost outcomes. Separate a market-price movement from a benefit created by negotiation, then distinguish proposed terms, implemented changes and savings that finance can validate.

For food / QSR	Why the buyer must check it
Price per case vs usable portion	Normalize pack weight, usable yield and waste; a cheaper case may be more expensive per served item.
Minimum order and shelf life	A larger discount can create expiry exposure or storage pressure; compare a feasible ordering pattern.
Product or supplier substitution	Keep specification, allergen, quality and applicable certification checks with the responsible specialists; do not let price override them.

EVIDENCE OF A BETTER WORKFLOW

Measure accepted comparisons, unresolved acknowledgments, repair time, supplier response quality and realized landed-cost benefit. Use the PPV convention on page 18 and the value bridge on page 45 consistently.

PRACTICAL WORKFLOWS

Planning and inventory

Use AI to explain exceptions; use validated models to calculate the response.

Oracle Planning, SAP Joule planning and Kinaxis Maestro offer enterprise planning capabilities. For an initial analytical test, use an approved export in Copilot in Excel or code-backed ChatGPT or Claude work to reconcile records, backtest assumptions and prepare scenarios. [\[S02, S11–S12, S42, S45–S46\]](#)

Step	What to do	Deliverable / control
1. Reconcile signals	Join store–SKU sales, stockouts, promotions, calendar, inventory, open supply and lead-time history.	A frozen input snapshot; item / recipe / unit mappings; missing-data register.
2. Test the baseline	Compare forecast with actuals by relevant horizon and segment; mark censored sales and one-off demand.	Bias and WAPE where the denominator is meaningful; no causal claim based only on a narrative.
3. Simulate alternatives	Call an approved forecasting or inventory model; test supply, capacity, shelf life, service and change costs.	Feasible options with assumptions; distinguish modeled supply from confirmed availability.
4. Review and implement	Planner accepts the trade-off; eligible execution uses the authorized system and is reconciled.	Decision version, approval and receipt; track service, waste, expedite cost and plan stability.

Food / QSR: connect menu demand to usable ingredients

For a QSR operation, start with store-level menu demand and translate it into ingredients using approved recipes, portion yields and pack conversions. Where data permit, include promotions, dayparts, menu availability and delivery cutoffs. Sales during a stockout understate demand. Check remaining shelf life at receipt and use, rather than relying on the expiry date alone.

Allocate eligible lots, using first-expire-first-out where the policy permits it. Protect the donating site's service before transferring stock, and respect handling and storage limits. Quality release and substitution approvals remain compulsory. This is a proposed control design; product-specific food-storage thresholds require the appropriate operating standards.

ORACLE COMPONENT LIMIT TO PRESERVE

The reviewed 26C Inventory Planning Command Center omits lead-time and upstream supply constraints from its recommendations. Check feasibility separately before releasing an action. This limitation belongs to that component and should not be generalized to all Oracle planning products. [\[S41\]](#)

A read-only pilot can replay a shortage against the response actually taken, comparing missed demand, waste, rework and expediting. The inventory model and its assumptions remain on page 21.

PRACTICAL WORKFLOWS

Logistics and import/export

Build an evidence-complete recovery or filing decision—not a confident status story.

1. From a late shipment to a confirmed recovery option

Where configured, Oracle OTM's Shipment Change Assistant supports shipment-related actions; project44 offers visibility and orchestration workflows to evaluate. An approved general-purpose workspace can assemble event and email evidence first, while a review flow routes material changes to the authorized owner. [S15, S38, S42, S45, S48]

Join shipment identifiers with event observation and receipt times, carrier acknowledgments, rates and delivery cutoffs. The exception queue should identify affected stores or customers, likely exposure, feasible responses, incremental cost and required authority. Distinguish the arrival estimate from a booking, confirmed capacity and an actual receipt throughout the case.

For food and QSR lanes, receiving windows, handling requirements, remaining shelf life and the stockout window determine whether a recovery is useful. A faster route may still arrive too late or breach product-handling conditions. Measure time to confirmed recovery together with expediting and waste costs.

2. From scattered documents to a reviewable shipment dossier

Use GTM and the authorized broker or Fasah workflow for their defined responsibilities. Approved ChatGPT or Claude analysis can extract and compare invoice, packing, transport and specification fields. Each material value should retain a locator back to the source document for review. [S05, S28–S29, S42, S45]

Check	Useful AI output	Human / system boundary
Goods and values	Differences in quantities, units, weights, price and named parties.	Resolve conflict with authoritative records; no invented missing fields.
HS, origin and permits	Candidate rationale and missing technical or origin evidence.	A qualified specialist validates classification, measures and applicable permits. [S27]
Submission and release	A completed, approved dossier plus status acknowledgments.	Only the permitted identity / channel submits; an agent does not grant clearance.

REGIONAL SCOPE CHECK

Check the requirements for the actual goods and jurisdiction. SABER should not be assumed to cover every food shipment; a certificate and a customs release have different functions. Permissions in one GCC system do not automatically extend to another. Route unresolved evidence to a specialist.

Track document completeness at first review, rejected filings, repair time and dwell exposure. Automated follow-up must identify the correct shipment and avoid duplicate messages or bookings.

CURRENT TOOLS + STARTER BRIEFS

The general-purpose AI workbench

Start with controlled analysis and preparation; keep commercial execution separate.

Tool / current scope	Practical use	Required check
ChatGPT Analyze uploaded tables with code; conduct source-linked research. [S42–S43]	PPV bridges, quote normalization and category evidence dossiers.	Request row-level provenance and reproducible calculations; review extraction and totals.
Claude Create / edit files and run analysis on supplied data. [S45]	Compare supplier documents; prepare RFQ packs and formula-based workbooks.	Inspect changed formulas and source pages; limit network and connector access.
Microsoft 365 Copilot in Excel Use editable Excel formulas, tables, PivotTables and charts. [S46]	Maintain a reviewable exception workbook inside an approved Microsoft environment.	Validate joins, units, formulas and refresh dates; workbook edits do not authorize a PO.
Gemini in Google Sheets Create tables, formulas, analysis and charts; reviewed page is Workspace Experiments. [S47]	Collaborative supplier trackers and scenario sheets in an eligible Workspace setup.	Check account availability; inspect source ranges and formulas before sharing.
Power Automate approvals Route requests for human sign-off; this is workflow automation, not AI reasoning. [S48]	Send a validated AI exception pack to the proper approver and record the response.	Preserve approval IDs and state; prevent duplicate retries; enforce ERP permissions separately.

Give the assistant a work brief, not a vague goal

Give the assistant the decision to support, the records and dates in scope, and the authoritative fields. Specify the output, source locators, reproducible calculations and treatment of missing data. State its action limits. Retain the inputs and accepted output for subsequent review.

Example — quote comparison: “Compare these approved quotations with the item master. Show pack conversions and source file, page and line. List missing freight or delivery terms. Produce a landed-cost comparison and questions requiring buyer judgment. Do not select a supplier or send a message.”

APPROVED WORKSPACE DOES NOT MEAN UNLIMITED DATA

Use only the employer-approved product, account and data scope. Check retention, connectors and permissions; omit unnecessary personal or commercial details. OpenAI states that business data is not used for training by default. This policy neither means zero retention nor grants permission to upload confidential records. [S44]

Before adopting the workflow, test a mismatched unit, missing quotation term, impossible lead time or instruction embedded in a supplier document. Check whether the assistant asks for evidence and respects its limits. A plausible-looking completion is not enough. Toolkit_Workflow_Briefs.md contains the complete starter briefs.

ANALYSIS

Procurement becomes policy-and-exception work

Buyers spend less time moving documents and more time deciding which exceptions deserve attention.

A procurement agent could follow a requirement from spend classification and specification retrieval through RFQ preparation, bid comparison, supplier-risk checks, negotiation and controlled purchase creation. Today's sourcing and supplier-communication documentation covers parts of that sequence. The practical task is to identify which parts work for the category and which still need buyer judgment. [S07, S14, S16]

How tasks change

Buyers and procurement specialists. Quote entry, comparison preparation and confirmation chasing are plausible areas for less manual work. The remaining queue will contain contradictory terms, missing responses, unusual demand and policy exceptions. Buyers need access to the original evidence so that they can resolve those cases, rather than approve a summary they cannot challenge.

Category and sourcing teams. Agents could bring overlooked transactions into review and surface smaller opportunities. Category owners would still decide the concentration they will accept, switching costs, negotiation principles and the supplier capabilities worth paying for. Price comparison alone cannot settle these choices.

Supplier management. A supplier with an unusual capacity problem or disputed term needs a named person to contact. Preserve that route when standard messages and negotiations become automated. A high volume of digital exchanges tells the business little about whether the relationship is working.

Automate earlier	Delegate after evidence	Retain judgment
Extract fields, compare like-for-like quotes, prepare follow-ups.	Release repeat approved orders; negotiate a bounded set of standard terms.	Choose critical suppliers; accept dependency; change technical specifications.
Flag duplicate spend and missing acknowledgments.	Handle low-risk exceptions using tested playbooks.	Resolve fraud concerns, supplier distress and conflicts between service and cost.

Define the task and its stopping conditions before delegating it. An instruction to maximize procurement savings leaves too much unresolved: the agent could lower unit price by accepting excess inventory, unreliable deliveries or unfavorable terms. A useful mandate states which outcomes must be protected and when the buyer must decide.

WRITE THE MANDATE

Specify the agent's objective, eligible work, permitted tools, spending authority, evidence requirements, escalation route, accountable owner and expiry date before it enters production.

ANALYSIS

Savings, PPV and strategic sourcing

Trace a proposed saving from the agreed baseline to the purchase, receipt and financial outcome.

Before reporting a saving, establish the baseline and how the benefit will reach the business. Track price variance, negotiated terms, implemented changes and finance-validated results separately. The agent can assemble the supporting records; category owners and finance should agree the measurement policy and decide whether the result meets it.

METRIC DESIGN

Illustrative PPV convention: $PPV = (\text{actual unit price} - \text{baseline unit price}) \times \text{actual quantity}$. Under this convention, a negative result is favorable. Publish the sign convention because companies use different ones.

What the bridge must isolate

A comparison needs equivalent specifications and units, with mix, volume, currency, commodity-index changes, rebates and freight shown separately. A falling commodity market can reduce prices without any negotiating gain. Larger minimum orders may add holding or obsolescence costs, while a conditional rebate has value only if the conditions are met and the rebate is earned.

For category decisions, connect commodity exposure and contract formulas to supplier capacity, geographic concentration and switching lead times. An agent can maintain this dossier and update the underlying evidence. Present a range of cost scenarios. For indexed contracts, first establish whether the agreed pricing formula was applied correctly before attributing a movement to negotiation.

Evidence layer	Agent contribution	Human / independent check
Spend normalization	Map supplier and item aliases; surface uncertainty.	Approve unresolved mapping and specification equivalence.
Should-cost scenario	Assemble input prices and contract assumptions.	Validate process knowledge, yield and commercial feasibility.
Risk-adjusted award	Compare landed cost, delivery and dependency.	Accept trade-offs and strategic relationships.
Realization	Reconcile PO, receipt, invoice, rebate and baseline.	Finance signs off on timing and double counting.

Procurement, planning, operations and finance need a common set of outcome definitions and a forum for settling disagreements. Otherwise each function can report success while transferring cost to another. The value review should show where price, service and working capital pull in different directions, rather than bury the trade-off in a single score.

Oracle provides source-to-settle and landed-cost capabilities relevant to this analysis. The value bridge proposed here is a measurement design, however, rather than an out-of-the-box savings promise. [\[S01, S05\]](#)

EVIDENCE + FORECAST

Can buyer agents negotiate with supplier agents?

The technical exchange is feasible in principle; agreeing what each agent may commit is harder.

Pactum describes policy-based procurement negotiation, and its Honeywell case reports more than 2,500 autonomous negotiations. This gives the discussion a practical reference point. It remains a vendor-reported deployment: the case does not provide a general savings rate or establish that supplier agents participated on the other side. [S16, S17]

Four different levels are often conflated

There are several distinct arrangements. In assisted negotiation, a person uses AI to prepare and compare terms. A buyer-side agent can instead bargain with a human supplier under a mandate. Bilateral agency gives both sides delegated authority. Network coordination goes further, adjusting interdependent capacity and commitments across several parties. Evidence for one arrangement should not be treated as evidence for all four.

Protocols can help agents discover each other and exchange tasks. Commercial agreements still need to establish who can bind each firm, which terms can change, what information can be shared and how disputes will be handled. UNCITRAL's 2024 Model Law on Automated Contracting offers legal context; applicability in any country requires a separate check. [S20, S26]

Design conditions for a credible bilateral pilot

A bilateral pilot should verify both identities and delegations, use a versioned term schema, set reservation limits and provide a human dispute route. Record offers, counteroffers and final acceptance as commercial states. Protect each party's confidential bids and reservation prices. The shared contract record needs explicit confirmation; private model reasoning need not be exchanged.

Buyer and supplier interests can still conflict when both use software. Faster bargaining could intensify price pressure, encourage misrepresentation or trigger repeated renegotiation. Multi-party allocation also raises questions about information sharing and competition law. These are risks to assess with counsel, not findings that the products reviewed here have caused such failures.

FORECAST

Standardized repeat purchases between trusted counterparties are the more plausible early setting for bilateral agents. Open autonomous bargaining across strategic supply networks remains speculative.

Evaluate the full distribution of negotiations, including nonresponse, rejected terms and failed execution after agreement. A favorable case study does not show how often the workflow produces a completed, compliant purchase.

ANALYSIS

Planning moves from periodic preparation to continual exception response

Frequent tactical replanning can support S&OP, while strategic commitments still need a human forum.

Forecasts, constraints and scenarios are already part of planning software. Oracle covers demand and supply planning; SAP describes exception analysis and simulation; Kinaxis combines AI with optimization and simulation. These computational capabilities provide the basis for the proposed agent workflow, rather than relying on a language model to construct a feasible plan unaided. [S02, S11, S12]

What changes in the demand-planning job

An agent could monitor signals, compare forecasts with orders and sell-through, retrieve promotion details and prepare scenarios. That would leave planners more time to assess assumptions, uncertainty and interventions. The forecast distribution should still come from a validated demand model, with its inputs and performance available for inspection.

The planner must determine whether a movement reflects lasting demand, a stockout, a one-off event, an assortment change or a data problem. Sales lost through unavailable stock are not evidence of weak demand. External explanations can help, but need corroboration before they become part of the planning assumptions.

What changes in S&OP / integrated business planning

Scenarios and routine reconciliations can be refreshed between meetings. Decisions about scarce capital, labor, supplier capacity, risk appetite and customer promises often need a periodic cross-functional forum. S&OP should use the more current evidence to decide these commitments, with less meeting time spent reconciling competing spreadsheets.

Continuous loop	Periodic human forum
Detect deviation; refresh scenarios; identify feasible corrective options.	Choose service priorities, capacity commitments and financial targets.
Reallocate within preapproved limits; monitor plan stability.	Approve material changes to network design, sourcing and customer promises.
Track assumptions and alert on invalidation.	Resolve competing functions' objectives and reset policies.

Frequent replanning has costs of its own: changed orders, disrupted schedules and uncertain supplier commitments. Build freeze windows, a cost of change and hysteresis into the policy. An agent should revise a plan when the likely benefit justifies the disruption and uncertainty, rather than react to every new signal.

THE PLANNER'S RESPONSIBILITY

Planners remain responsible for the assumptions, decision quality and commitments behind the plan, even as forecast editing and scenario preparation become less manual.

ORIGINAL MODEL + ANALYSIS

Inventory autonomy needs math and business constraints

A forecast is only one input to a replenishment decision.

A replenishment workflow needs to revisit demand and lead-time uncertainty, available supply, commitments and service policy before proposing new orders. Inventory-oriented agent workspaces point in this direction. Their usefulness still depends on accurate records and an appropriate model for the items being managed. [S06, S09]

MODEL – ASSUMPTIONS MATTER

Illustrative continuous-review model: reorder point = expected demand during lead time + safety stock. With independent daily demand and lead-time variability: $SS = z \times \sqrt{\text{mean } L \times \text{variance } d + \text{mean } d^2 \times \text{variance } L}$.

Assume independent daily demand with mean 100 units and standard deviation 20; mean lead time 10 days and lead-time standard deviation 2 days. Using $z = 1.645$ gives safety stock of about **345 units** and a reorder point of about **1,345 units**. Ignoring lead-time variability would reduce the calculated safety stock to about 104 units—a very different decision. These are synthetic values, not a recommended target.

The approximation assumes stationary demand, independence between daily demand and lead time, and a normal service approximation. Intermittency, promotions, correlated disruptions, perishability, substitutions and multiple inventory echelons can require other models or simulation. A 95% cycle-service assumption does not guarantee 95% fill rate. Periodic-review systems must include demand over the review period as well.

What the agent manages—and what people decide

Agents can identify unusual lead-time variation, suspect master data and target changes, then call an inventory optimizer and release eligible actions after checks. People set service differentiation, acceptable shortage exposure, supplier constraints and obsolescence policy. Better forecasts will not necessarily reduce stock when the service promise or supply uncertainty has increased.

Base the decision on a defined inventory position that includes relevant open supply, reservations and backorders. Respect minimum orders, case packs, shelf life, storage capacity and approved sources. Separate quarantined stock and stock belonging to other entities from usable supply; the on-hand balance alone does not show what can serve the next demand window.

PARAMETER CONTROL

Version inventory targets with their assumptions and review conditions. Assess the exposure before applying changes across large numbers of reorder points.

PRODUCT BOUNDARY

Oracle's 26C command-center limitations are set out on page 44. Check suggested replenishment against the full supply constraints before release. [S41]

ILLUSTRATIVE FUTURE OPERATING SCENARIO

Six agents, one disruption, one accountable owner

A shared food/QSR case carries the same shortage through analysis, commercial action and receipt.

Consider a possible 2030–2033 workflow. A chilled-ingredient shipment is delayed by 36 hours before a QSR weekend peak. The team has to cover Friday's service window using the approved product and eligible stock. This is an illustrative operating scenario, not a vendor forecast.

Agent / handoff	Contribution to the same versioned case
01 Demand agent	Re-estimates the protected Friday demand window at 1,200 cases from orders, promotions and stockout-adjusted sales.
02 Inventory agent	Confirms 700 usable cases locally and 200 transferable cases after protecting the donor site. Excludes 150 quarantined cases; checks lot life.
03 Planning agent	Tests arrival cutoffs, storage and shelf life. The selected plan needs 300 additional cases; compares transfer, approved purchase and service restriction.
04 Procurement agent	Requests 300 cases of the same approved SKU under existing terms. Checks total exposure; releases the PO within its mandate only after supplier acceptance and policy checks.
05 Supplier agent	Reserves 300 eligible cases and returns lot / expiry evidence, price, quantity and promised dispatch in an authenticated acceptance.
06 Logistics agent	Confirms carrier capacity and receiving slots for the transfer and purchase. On receipt, reconciles quantity, condition, lot and release with ERP / WMS.

PLANNED COVERAGE: $700 + 200 + 300 = 1,200$ CASES

The planned 1,200 cases comprise 700 local, 200 transferred and 300 from approved supply. Neither the delayed inbound nor the 150 quarantined cases is available for that window. Keep the case open until receipt, quality release and store allocation records confirm that the plan was delivered.

Where humans and hard controls enter

One orchestrator maintains the case ID and input version. A change in price, SKU, lot eligibility or arrival time triggers renewed checks. Missing temperature or quality evidence, an unapproved substitute, or excess exposure goes to the accountable person. Without supplier or carrier confirmation, the team must replan rather than mark the shortage resolved.

The sequence is detect → diagnose → simulate → authorize → execute → reconcile. Investigate the cause separately from the observed delay, retaining uncertainty where evidence is missing. The supplier-agent handoff assumes trusted identities, shared terms and working integrations. Planning, logistics and protocol sources support individual components of this design; they do not document this complete deployment. [\[S02, S15, S20–S21, S25\]](#)

FORECAST + DESIGN

Dashboards remain—but the operating interface changes

Keep a common performance record, with conversation for investigation and an action queue for follow-through.

The proposed interface has three parts: a shared scorecard, a conversational investigation layer and an agent-managed action queue. Leaders still need comparable measures over time. Conversation helps someone ask why a metric moved, inspect evidence or test an alternative; the scorecard and case records retain the agreed measures and commitments after that exchange ends.

What a useful analytics agent should do

Take the question, 'Why did inventory rise?' A useful answer names the population and dates, applies the agreed inventory definition and reconciles receipts, demand changes and blocked stock. Where possible, it separates price, volume and mix. The explanation should link to records and identify incomplete coverage so that the proposed cause can be tested.

A control tower should put the most urgent business decisions in front of the team. For each exception, show the service or cash exposure, available responses, approval needed and age of the evidence. A small deviation can threaten a critical item, while a large financial anomaly may be less urgent. Safety and compliance need their own priority rules.

Question	Weak response	Decision-grade response
Why is PPV adverse?	A narrative based on supplier names.	A reconciled bridge for baseline, price, mix, FX and quantity.
Will we stock out?	A single forecast or red KPI.	A probability / range, assumptions, affected customers and options.
What did the agent fix?	Number of messages or tool calls.	Verified outcome, cost, remaining risk and any human rework.

This requires agreed meanings for service level, savings, inventory days and lead time, including the denominator, timing convention and owner. Changing those definitions from one conversation to the next makes the answers incomparable. Totals should reconcile to the operational data or ledger that finance and operations use for the same measure.

CONTROL-TOWER DESIGN

Retain the scorecard as the common record. Use conversation to investigate it and agents to track authorized actions through to a verified result.

This is the report's proposed interface design. The reviewed evidence does not establish that any vendor has removed the need for conventional reporting.

PROPOSED KPI DICTIONARY

New KPIs: measure safe outcomes, not agent activity

A proposed measurement dictionary for a hybrid human-agent operation.

Procedure-aware evaluation research shows that an agent can appear to finish a task while violating required steps. Its benchmark is not a live supply-chain error rate. It does support testing both the operational result and compliance with the required process. [\[S22\]](#)

KPI / definition	Interpretation / guardrail
Compliant completion = verified outcomes with all required controls passed ÷ eligible attempts.	Use a fixed eligibility policy. Include failed and abandoned eligible attempts in the denominator.
Autonomous completion = compliant outcomes needing no human action ÷ eligible attempts.	Separate preauthorized policy from case-by-case approval; do not count drafts as completed work.
Human intervention burden = review + repair minutes ÷ eligible attempts.	Measure burden, not simply the fraction of cases touched. Include downstream rework.
Decision latency = verified trigger to authorized executable decision.	Also track execution latency separately; a carrier delay is not all agent reasoning time.
Unauthorized-action rate = actions outside valid authority ÷ attempted actions.	Use zero tolerance for material breaches; independently inspect denied attempts and near misses.
Reconciliation failure = committed actions lacking matching verified state ÷ committed actions.	Set time windows by process; detect duplicates, lost updates and unacknowledged commitments.
Net realized value = validated benefit – incremental operating / loss costs.	Separate recurring P&L , one-time cash release and redeployable labor capacity.
Plan stability = material revisions inside a defined freeze window.	Pair with service and shortage risk; stability alone can reward failure to respond.
Calibration / freshness = forecast reliability and input-age compliance.	Use suitable probabilistic calibration tests; an LLM confidence label is not calibrated risk.

Assign an owner, event source, reporting frequency and exclusion rule to every metric. Procurement operations can own completion and rework; finance validates realized value; risk and IT independently inspect authority and incidents. Pair these measures with fill rate, inventory days, lead-time variation and supplier outcomes to keep the business effect visible.

INTERPRETING ESCALATION

A correct escalation can lower the touchless-completion rate while preventing a loss. Read autonomy measures alongside risk, rework and verified business outcomes.

ANALYSIS

Logistics, HS classification and import/export

Document work can move faster while classification, filing and clearance retain their separate responsibilities.

The Harmonized System uses a six-digit nomenclature and classification rules, with national extensions and measures requiring local checks. A brief product description may not settle the classification. The dossier should establish material, function and composition and retain the technical documents needed to support the proposed code. [S27]

A practical trade-agent division of labor

Begin by comparing the invoice, packing, transport and goods records: quantities, values, origin and parties should agree. Retrieve the relevant classification and regulatory material with its date, then prepare the dossier or draft declaration for review. Submission is a separate step, performed only through the permitted system channel and an authorized broker or company identity.

Classification, origin and valuation answer different questions. Even a well-supported HS code does not establish preferential origin, admissibility, duty treatment or permit compliance. Ask for the missing evidence; a favorable inference made merely to speed clearance can leave the substantive question unresolved.

Logistics agents can assemble rate comparisons, request confirmations, prepare claims and monitor missed milestones. Changing a booking, customer promise or insurance exposure requires separate authority. The transport and trade platforms reviewed here supply useful tools, but the implementation must define who can act on their information and how the resulting commitment is confirmed. [S04, S05, S15]

Earlier automation	Conditional autonomy	Human / specialist boundary
Document extraction, cross-document checks, missing-field requests.	Repeat filing preparation from validated master data.	Ambiguous classification, contested origin or valuation.
ETA monitoring and approved-rate comparison.	Rebooking within approved carriers, lanes and costs.	Unusual liability, regulated cargo or material customer impact.
Permit-expiry and screening alerts.	Routine updates through authorized integrations.	Potential sanctions match, legal exception or disputed declaration.

CLASSIFICATION RECORD

A similar shipment is a useful starting point, not sufficient classification evidence. Retain the applicable technical basis, source date and approval for each material trade decision.

These are process-control recommendations. Applicable filing rights, liability and regulatory interpretations should be resolved by a qualified trade or legal specialist for the jurisdiction concerned.

OFFICIAL EVIDENCE + ANALYSIS

Saudi Arabia: existing digital rails matter

Use the documented trade services as the starting point for enterprise coordination.

ZATCA's 'Provide an Issued Declaration' service describes a broker creating a statistical declaration in Fasah from export-manifest data supplied by the shipping agent, with AI-based validation afterward. This is evidence for a defined service. It does not describe all declaration processing or establish autonomous customs clearance. [S28]

SABER supports product and shipment conformity services; conformity and customs clearance remain distinct. ZATCA also documents GCC-related electronic declaration coordination. The earlier SABER–Fasah linkage statement was not revalidated and should retain the qualification recorded in the source register. [S29, S30]

Verified foundation	Potential enterprise use	Boundary
Fasah declaration workflow	Prepare consistent goods and shipment evidence; reduce repeated entry.	Government/broker authorization and service-specific scope. [S28, S29]
SABER conformity platform	Track product evidence, certificates and shipment readiness.	Do not treat a certificate as blanket clearance permission. [S30]
National logistics / industrial strategy	Align enterprise data, skills and logistics transformation with documented priorities.	The cited NIDLP plan is historical, covering 2021–2025. [S32]
Personal-data transfer regulation	Map personal data, processors and transfers in the agent architecture.	Applicable conditions and safeguards require case-specific review. [S31]

What is credible for Saudi organizations now

A practical first step is a shipment dossier connecting the supplier, item, specifications, contract, invoice and required evidence. An internal agent can identify omissions, request documents and route discrepancies. Move from that preparation to permitted integrations that return acknowledgments, rather than using screen automation to work around portal access controls.

Test Arabic and English records together, including supplier aliases, units and legal-entity identifiers. Preserve discrepancies for review. A translated certificate or specification still needs to be checked for meaning and completeness; translation alone does not establish equivalence for the relevant process.

REGIONAL STARTING POINT

A ready-to-file, approved dossier is a useful initial outcome. It shortens preparation while keeping the authority to submit and clear the shipment with the parties responsible for those decisions.

ANALYSIS

A future GCC supply network

Interconnection must respect separate systems, mandates and commercial interests.



Conceptual future network — no shared sovereign or commercial authority implied

The proposed network links enterprise ERP and finance to suppliers, forwarders, carriers, ports, warehouses, conformity services and customs workflows. Each participant remains the source of authority for its own events and commitments. Oman's customs portal and Bayan services illustrate the separate national systems that regional coordination has to work with. [S33]

A realistic sequence of regional integration

Start by aligning identifiers and exchanging permitted documents, evidence and status acknowledgments. Known counterparties can then test approved booking or replenishment changes. A later step would share selected capacity or service signals to coordinate conditional commitments. That last step is a forecast: the evidence reviewed does not describe a fully integrated autonomous GCC network.

At a terminal, the proposed digital action must meet a physical slot, handling capability and accepted handoff. An arrival estimate alone cannot confirm those conditions. This research does not establish how widely Saudi ports currently use AI, and national logistics ambitions should not be substituted for deployment evidence.

Data governance is not a localization slogan

SDAIA's personal-data transfer regulation specifies conditions and safeguards, with risk assessment in defined circumstances. It does not support a blanket statement that all supply-chain data must stay in Saudi Arabia. Legal and security teams should classify the data and assess processing, hosting, access and transfers against the requirements that apply. [S31]

FORECAST

The planning case is a federation of trusted, permissioned workflows. Suppliers, ports, banks and customs retain separate commercial or sovereign authority rather than reporting to one controlling agent.

ANALYSIS

Data and interoperability determine the ceiling

A connected system can still disagree about what the same message means.

Connecting a network requires agreement at four levels: the systems must exchange messages, participants must interpret them consistently, operational states must be acknowledged, and commercial rights must be understood. A technically successful connection can still fail at any of the other three levels, leaving a commitment ambiguous.

MCP provides interfaces to tools and data; A2A addresses agent discovery and task exchange; GS1 EPCIS defines visibility-event sharing within and between firms. They help at different layers of the connection. Purchasing authority and confirmation that a transaction completed must still be supplied by the business process. [S20, S21, S25]

Minimum data contract for an autonomous loop

Object / control	Required question
Identity	Are supplier, legal entity, item and location identifiers unique and mapped?
Meaning	Are unit, currency, Incoterm/version, ownership and event state explicit?
Freshness	When was the fact observed, received and last verified? Is it stale for this decision?
Authority	Which system owns the field, and who may change or dispute it?
Transaction state	Is this a proposal, approval, commitment, acknowledgment or completion?
Recovery	Can a replay create duplication? Can the action be compensated or safely halted?

Measure master-data quality for the decisions being delegated. A high overall completeness score can hide the missing specification that prevents one shipment from being classified. Check critical fields, exceptions and lineage in the eligible population. Unresolved records belong in a review queue, not in a workflow that fills the gaps with plausible guesses.

Cross-company execution needs explicit states and correlation IDs. A partially completed workflow may require compensating actions, since it cannot always be committed or reversed as one transaction. Specify whether the agent can cancel a booking itself or merely request cancellation, and retain any obligation that remains afterward.

MASTER-DATA CHANGES

Treat a proposed master-data correction as a separate governed change, with evidence, an owner and an impact check. Do not bury it inside routine task execution.

ANALYSIS

Cybersecurity and segregation of duties

Keep permissions narrow, especially when an agent reads material from outside the organization.

A supplier email, attachment or web page may contain instructions intended to redirect the agent, expose records or trigger an unauthorized action. Treat external material as evidence to inspect, never as authority to change permissions or policy. The controls below are proposed safeguards informed by lifecycle risk guidance. [\[S19\]](#)

Risk scenario	Preventive control	Detection / recovery
Invoice or email tells agent to change bank details.	Separate bank-master workflow; independent verification; no payment privileges.	Alert on attempted changes; freeze affected payment path.
Agent exceeds a purchasing limit through split orders.	Aggregate exposure checks by supplier/category/window, not only per line.	Audit cumulative commitments and related transactions.
Tool timeout causes duplicate order or booking.	Idempotency key, state readback and bounded retry.	Reconcile duplicate commitments; initiate controlled reversal.
Sensitive bids or personal data leak to a counterparty.	Least disclosure; approved data destinations; role-scoped retrieval.	Log egress; revoke identity; investigate affected records.
Policy or model update changes behavior silently.	Versioning, regression tests, staged release and rollback.	Compare safe completion, overrides and anomalous tool use.

Separate design, approval, execution and assurance

Give each agent its own identity and only the privileges needed for its work. Configuring the agent should not also grant the right to approve its purchases or suppress its audit trail. Microsoft's procurement-agent identity and Oracle's role requirements show that these are concrete implementation choices. [\[S07, S13\]](#)

Enforce budgets, eligibility, schemas, access rights and approval gates outside the language model wherever possible. An instruction to follow policy cannot block a transaction on its own. A second model may help screen unusual cases, but it should not be the only check on the first model's errors.

TEST THE STOP PROCEDURE

Before expansion, demonstrate that revoking the agent's authority stops new commitments, preserves the investigation record and leaves a usable manual fallback.

ANALYSIS

Accountability: an agent is not the accountable executive

Name the people responsible for the objective, delegated exposure, system operation and independent review.

Every autonomous workflow needs a named human owner. The company remains responsible for its governance choices, while liability between it, the software provider, service providers and counterparties depends on the applicable contracts and law. Automated-contracting frameworks provide context rather than a universal allocation of responsibility; counsel must determine the obligations that apply. [S26]

Decision / role	Recommended accountability
Business process owner	Own objective, eligibility, control design, monitoring and service outcome.
Delegated approver	Accept the specific exposure or policy envelope within corporate authority.
Technology / model owner	Own integrations, identity, versioning, testing, availability and rollback.
Data owner	Own critical definitions, quality remediation and authoritative records.
Risk / legal / compliance	Interpret obligations; challenge controls; approve relevant exception policy.
Internal assurance / finance	Independently inspect controls and validate realized business benefits.

A practical authority ladder

Execution rights can be granted in stages: read approved records; recommend a decision; prepare a nonbinding draft; then commit within a defined policy. Changing that policy is a separate authorization. A history of successful cases should not silently give an agent broader discretion than the organization originally approved.

Reviewers need the evidence, time and expertise to challenge a decision. Thousands of routine approval clicks can produce little real oversight and encourage automation bias. Design the queue around material changes and unusual exposure, so that human review is concentrated where judgment can alter the outcome.

When an agent makes an incorrect decision

After an incorrect action, contain further activity and preserve the case and transaction records. Identify outstanding commitments, notify the owner and affected functions, then repair through authorized channels. Establish whether the cause was a poor recommendation, failed control, execution defect or external event; each requires a different corrective response.

ACCOUNTABILITY RECORD

Agree the agent mandate and responsibility map before deployment. Retain the audit trail needed to explain who authorized the action and what was executed.

ANALYSIS

The hybrid supply-chain workforce

Automation changes tasks within roles; staffing effects depend on how the business uses the released capacity.

The ILO's 2025 analysis distinguishes exposure to generative AI from job disappearance. In a separate field study, 5,172 customer-support agents achieved an average 15% productivity gain with AI assistance, with differing effects across workers. These findings inform the workforce discussion, but neither estimates procurement redundancies or autonomous supply-chain staffing. [S23, S24]

Role	Less manual work	More valuable human contribution
Buyer / procurement specialist	Quote handling, confirmations, routine comparisons.	Exception judgment, policy design and supplier problem solving.
Demand / supply planner	Data assembly, scenario preparation, routine parameter review.	Uncertainty assessment, constraints and business trade-offs.
Strategic sourcing	Research collation, spend cleaning and routine negotiation preparation.	Category strategy, market structure and durable supplier relationships.
Logistics team	Tracking requests, document reconciliation, standard rescheduling.	Operational recovery, service commitments and complex claims.
Import / export specialist	Document extraction and completeness checks.	Classification evidence, regulatory interpretation and authorized filings.
Supply-chain leadership	Status compilation and fragmented escalations.	Decision rights, resilience, organizational design and investment choices.

Staffing may still change. With fixed demand and service scope, sustained net time savings could reduce hiring or transactional positions. A growing business might instead use the capacity for wider spend coverage, stronger controls and supplier development. Management choices and implementation economics will determine which path is taken; exposure alone cannot predict it.

Manage agents as operating resources with named identities, scopes, budgets, owners, quality measures and retirement rules. A person may supervise several workflows, but the appropriate workload depends on the severity and frequency of exceptions. An arbitrary agent-to-person ratio would miss the work needed to review and repair failures.

KEEP THE LEARNING PATH

Replace the learning lost when junior tasks are automated. New staff still need supplier contact, data investigation and commercial decisions from which to develop judgment.

The durable skill is knowing how to frame a decision, examine its evidence and commercial consequences, and improve the process behind it. Prompt-writing is only part of that work.

ILLUSTRATIVE – SYNTHETIC

Economics: separate capacity, profit and cash

Synthetic assumptions show why staff capacity, recurring benefit and one-time cash release need separate treatment.

Illustrative assumption	Calculation	Result
24,000 cases/year; 12 minutes saved/case	$24,000 \times 12 \div 60$	4,800 gross hours
1,200 hours/year new review and operating work	$4,800 - 1,200$	3,600 net hours
1,800 productive hours per FTE/year	$3,600 \div 1,800$	2.0 FTE capacity
Realized benefit on SAR 40m eligible spend: 0.6%	$40m \times 0.006$	SAR 240k/year
Average inventory falls from SAR 12m to SAR 11m	$12m - 11m$	SAR 1m one-time cash release
Marginal carrying-cost rate: 12%	$1m \times 0.12$	SAR 120k/year modeled cost benefit
Incremental annual operating cost: SAR 300k	$240k + 120k - 300k$	SAR 60k recurring net benefit

The base result excludes implementation costs and any cash value for staff capacity. Releasing two FTE of time does not automatically remove two jobs or save their salaries. The carrying-cost rate must reflect avoidable costs without duplicating another finance benefit. The SAR 1m inventory reduction is a one-time balance-sheet release, not annual profit.

Sensitivity matters more than a polished headline

At 0.3% realized procurement benefit and a SAR 0.3m inventory reduction, the same model gives SAR 120k + SAR 36k – SAR 300k = a SAR 144k annual loss. At 0.9% and SAR 1.5m less inventory, it gives SAR 360k + SAR 180k – SAR 300k = a SAR 240k annual gain. No probabilities are assigned to these assumed cases.

An investment appraisal must also include implementation, integration, evaluation, change management, supplier participation and incident losses. Identify incremental cash flows and benefits that depend on adoption. For a multi-year case, apply the organization's approved discount rate and distinguish the ramp-up period from steady-state operation before interpreting the return.

BUILD THE RETURN FROM EVIDENCE

Use measured eligible volumes and verified results to build the business case. The available evidence supports testing productivity gains, not promising uniform reductions in stock, lead times or staffing.

Track service, stockouts, waste, complaints and compliance alongside savings. Otherwise, a favorable procurement result may simply move cost or risk to another part of the business.

ANALYSIS

Counterarguments and failure modes

The right level of autonomy depends on the decision, its risks and the cost of controlling it.

Assumption to challenge	Counterargument	What would resolve it
Better models make integration unimportant.	Execution still depends on correct identities, current facts and valid transaction rights.	Demonstrate reliable state reconciliation under faults.
Multi-agent systems are inherently superior.	Extra coordination can add latency, inconsistency and attack surface.	Beat a simpler workflow on compliant outcomes and full cost.
Autonomy always reduces inventory.	Higher service targets or uncertainty can justify more inventory.	Measure risk-adjusted service and working capital together.
Human approval guarantees safety.	Overloaded reviewers may rubber-stamp plausible recommendations.	Test reviewer understanding and intervention effectiveness.
A shared protocol creates a shared network.	Counterparties may reject data sharing or have conflicting incentives.	Confirm participation, semantics and enforceable commitments.
A successful pilot can be copied everywhere.	Tail spend differs from direct materials; one lane differs from a global network.	Repeat tests across materially different risk and data segments.

When simpler automation is preferable

Stable inputs and known exception rules often call for deterministic automation. Constrained allocation calls for optimization. Language models are most useful where interpreting documents, clarifying requests or coordinating work adds enough value to cover the uncertainty and operating cost. Engineering guidance supports starting with a simpler design and adding agent complexity when the task justifies it. [\[S18\]](#)

Correlated reactions are a network-level risk. If participants respond to similar signals and objectives, they could all expedite, cancel or substitute together. This is a plausible failure scenario, not an industry-wide effect measured here. Exposure limits, varied evidence, damped plan changes and human coordination during disruption can help contain it.

COMPARE THE ALTERNATIVES

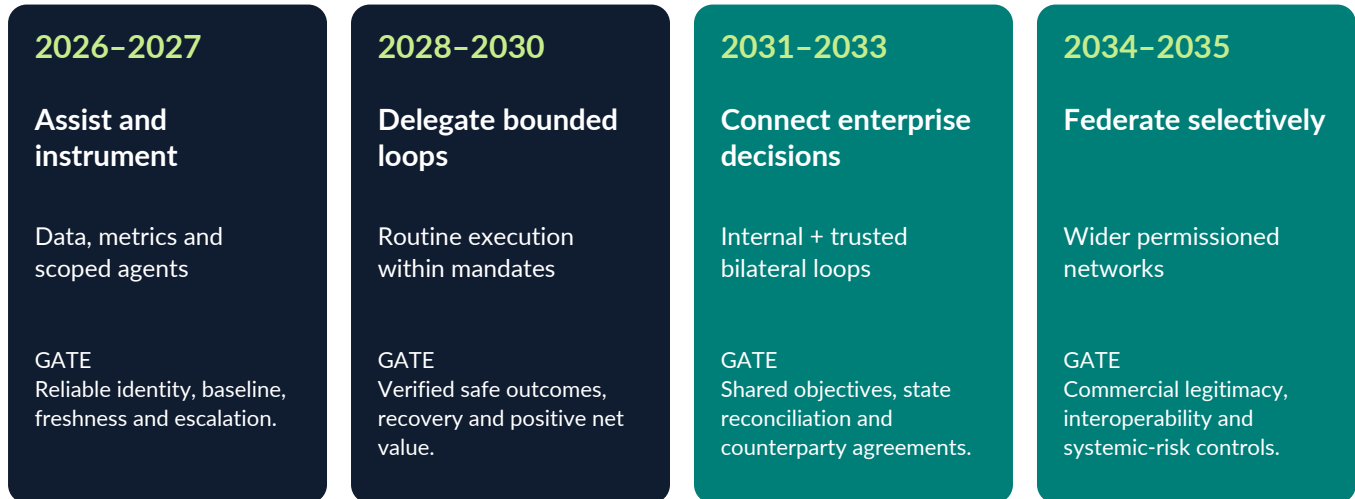
Choose the process that delivers the better controlled outcome. A dependable assisted workflow may justify investment more readily than a fragile autonomous one.

Adoption can stall even when the technology works. Participants must agree on definitions, permitted information sharing and responsibility for losses before coordinated decisions become acceptable business practice.

FORECAST – CONDITIONAL

2026–2035 roadmap: decision rights expand conditionally

These windows organize preparation; the milestones depend on readiness rather than a promised release date.



Constrained adoption / Selective autonomy / Accelerated federation

2026–2027. Establish governed data, shared metrics and narrowly scoped agents. Confirmation follow-up, evidence assembly and purchase preparation are useful early candidates. Expansion requires verified identity, current inputs, baseline performance measures and an exception queue that the operating team can manage.

2028–2030. Give proven workflows limited execution rights. Repeat buying, approved rescheduling and parameter updates become candidates where eligibility and recovery are clear. The meaningful change is organizational permission to commit a routine action without individual approval, supported by evidence that the full process works.

2031–2033. Connect selected procurement, planning and logistics decisions internally, then test bilateral commitments with trusted partners. Shared objectives become essential: a purchasing agent should not report success while its decision silently worsens another function's service or risk.

2034–2035. Permissioned networks may coordinate more broadly in standardized, well-observed segments. Open-network autonomy, unrestricted strategy changes and the removal of all human judgment remain speculative. Under the constrained scenario, assisted and agent-managed workflows could still dominate at the end of the period.

FORECAST

Some low-risk workflows may progress earlier. Others may retain individual approval indefinitely because their exposure or uncertainty makes that the better operating choice.

Progress would be helped by more reliable models, affordable evaluation, stable integrations, common event meanings and supplier participation. Fragmented master data, unresolved liability, security incidents, weak returns and unchanged approval structures could delay it. The roadmap is this report's conditional synthesis of the capabilities and constraints reviewed.

ORIGINAL SCENARIOS

Three scenarios and the gates that distinguish them

The scenarios differ in readiness and coordination; the evidence does not justify numerical probabilities.

Scenario	2026–2030	2031–2035
Constrained adoption	Assistants and selected agent workflows expand, but approval queues and data problems persist.	Autonomy remains narrow; networks mostly exchange visibility and documents.
Selective autonomy — planning case	Stable loops gain execution rights after demonstrated controls and value.	Cross-functional orchestration and trusted bilateral networks spread unevenly.
Accelerated federation	Reliability, interfaces and supplier participation improve together; economics are favorable.	Broader network coordination in standardized segments; strategic and legal boundaries remain.

Observable inflection points

The first inflection is an explicit execution mandate backed by independent checks and reconciled transactions. The next is cross-functional agreement on outcomes, with agents using the same authoritative state. Network delegation requires an additional agreement: counterparties must accept the identities, term definitions, conditional commitments and dispute arrangements on which the exchange depends.

Gate	Evidence to require before expanding authority
Reliability	Representative replay and live evidence, including rare severe exceptions and false success.
Economics	Net realized value after integration, review, inference and incident costs.
Governance	Signed owner, permission map, segregation of duties and tested revocation.
Network readiness	Counterparty participation, acknowledgments, data-sharing agreement and recovery path.

Track the share of eligible transactions completed safely, repair minutes, policy breaches, time to reconcile external commitments and reliable event coverage among critical counterparties. These measures show whether the operating model is improving before the effect is visible in enterprise financial results. Counting product announcements does not answer that question.

WHEN TO PAUSE

Narrow or stop a workflow that repeatedly exceeds its mandate, leaves material evidence gaps, loses transaction state or delivers negative net value after stabilization. A faster model does not resolve those failures by itself.

The scenarios are preparation tools, not predictions that a particular company, country or vendor will reach a specified stage by a fixed date.

ANALYSIS

An organizational action plan

Use a small, measurable scope to decide whether the organization should delegate more work.

Window	Management action	Required output
Days 1–30	Choose one procurement and one planning loop; baseline work and failure costs.	Decision inventory, eligible population, metric definitions and named owner.
Days 31–60	Implement read-only / shadow operation; repair critical data; test bad cases.	Evidence-backed recommendations, error taxonomy and authority design.
Days 61–90	Run approved execution on a restricted cohort; verify outcomes independently.	Measured net value, safe completion, review burden and a go / hold decision.
Months 4–12	Expand only proven segments; integrate downstream reconciliation and supplier feedback.	Operating playbooks, workforce plan, incident drills and benefit validation.

Design the pilot to answer a decision

Choose a question the pilot can answer. Does confirmation follow-up reduce unresolved acknowledgments without more complaints or incorrect PO changes? Does a shortage playbook improve service at an acceptable stock and expediting cost? Define the population and comparison before starting; an open-ended AI transformation exercise cannot produce a clear investment decision.

A central team can provide identity, integration, testing, security and shared metrics. Business ownership should remain with the function responsible for the outcome. That arrangement brings technical capability into the process without moving commercial decisions to an AI team that does not carry their consequences.

Buy, configure or build?

Test released native features first when they inherit the right records and controls. A specialist tool is worth considering for a measured gap. Custom orchestration is appropriate when the process is genuinely distinctive and the business can maintain it. Compare lifecycle cost and the ability to exit, not just the license or demonstration.

Review contract terms for data use, sub-processors, retention, logs, incident response, model changes, exportability, rate limits, support and failed actions. Plan the operational exit as carefully as deployment. Revoking an agent's rights should still leave the company able to buy, plan and reconcile its core transactions.

STEERING-COMMITTEE APPROVAL

Approve the pilot's scope, controls, owner and evidence threshold. Use its results to decide whether broader delegation is justified.

The schedule is a proposed sequence. It does not promise that any particular pilot will be ready for production within 90 days.

ANALYSIS

A professional development agenda

Develop commercial judgment together with the analytical and control skills needed to supervise automated decisions.

Capability	Practical evidence of competence
Supply-chain fundamentals	Explain service / cost / cash trade-offs, lead-time uncertainty and supplier dependency.
Analytical fluency	Reconcile PPV; distinguish forecasts from causal claims; understand distributions and constraints.
Data / integration literacy	Trace a purchase or shipment across systems; inspect identifiers, units, lineage and stale records.
Agent operations	Write an executable mandate, test normal and failure cases, interpret logs and manage escalation.
Governance / commercial judgment	Identify authority, conflicts of interest, sensitive data and material contract exceptions.
Human coordination	Resolve ambiguous requirements, negotiate difficult trade-offs and maintain supplier trust.

For procurement specialists and buyers

Build a portfolio case from imperfect spend or quotation records through to a purchasing recommendation. Show how you normalized the bids, handled uncertainty, calculated landed cost and checked authority and realization. A reviewer should be able to trace the decision, including the assumptions you rejected or left unresolved.

For planners and logistics professionals

Learn to interpret forecast distributions and scenario constraints, and distinguish predicted events from confirmed ones. Develop an exception playbook with a case where the cheapest or fastest response would fail operationally. Be able to explain the evidence that would justify action and the conditions that require escalation.

For managers

Supervising automated work requires clear objectives, a realistic review workload and an escalation route people trust. Challenge apparent success and ensure new entrants still gain experience with suppliers, troubleshooting and data stewardship. They need to learn how decisions are made, not only how to approve the recommendations placed before them.

A PRACTICAL COMPETENCE TEST

**Explain the proposed action, its evidence, the authority required and the way its outcome will be checked.
A professional should be able to defend all four.**

Much manual coordination could be displaced by this operating model. Some tasks or positions may shrink, while the consequences of poor judgment grow as execution becomes faster. The report therefore emphasizes commercial and operational understanding; it does not promise a particular employment outcome for any role.

ANALYSIS

Conclusions: the operating model is the strategy

The strategic choice is which decisions to delegate, under what controls, and with whose agreement.

ERP already provides workflows, analytical tools and selected agents. The next layer is likely to coordinate decisions across those systems. Replacing the application estate with a general-purpose model would discard established records and controls without resolving the execution problems identified in this review.

Specific, bounded forms of autonomy have support in product documentation. Evidence for sustained, independently verified enterprise-wide outcomes is thinner, and open multi-party networks are further from that standard. Investment should progress as individual workflows demonstrate reliable results, rather than assuming that a successful pilot validates every category or site.

Human responsibility becomes more consequential as routine work is delegated. Strategic commitments, regulatory interpretation, scarce-resource allocation and supplier relationships involve objectives and obligations that someone must own. Approval arrangements need to give that person a genuine opportunity to examine the evidence and alter the proposed decision.

The potential gain lies in fewer broken handoffs and faster decisions that hold up in execution. Isolated improvements in forecasting or RFQ volume may contribute, but can be outweighed by costs shifted elsewhere. Shared outcome definitions are needed to reveal when one function's apparent gain creates another's shortage, excess stock or risk.

Saudi and GCC organizations can build on existing trade, conformity and customs services. Data protection, permitted access and separate national authorities remain part of the design. A proposed future network should describe how these participants cooperate, while keeping its forecast distinct from the capabilities documented in current government services.

Five questions for the next leadership meeting

Question	Decision it forces
Which loop is worth closing first?	A measurable scope rather than an enterprise-wide slogan.
What may the agent commit?	Explicit delegation and protected boundaries.
What evidence would stop us?	A credible risk and value threshold.
Who benefits—and who carries the cost?	An enterprise outcome rather than departmental savings.
What must remain human?	Deliberate judgment, accountability and skills development.

THE OVERALL JUDGMENT

AI is likely to displace a substantial amount of manual supply-chain coordination. Its effect on particular jobs will depend on how tasks, investment and released capacity are managed, rather than one technical milestone.

EVIDENCE-ANCHORED FRAMEWORK

Function transformation matrix • 1 of 2

The autonomous column is a conditional target state—not a statement of current availability.

Function	Today	AI-assisted	Agentic	Autonomous	Future human role
Spend analysis	Manual / BI classification and reconciliation.	Suggest mappings; explain variance.	Continuously detect and investigate opportunities.	Publish validated recurring analyses under fixed definitions.	Own taxonomy, baseline and finance validation.
Requisition & PO	Workflow approvals and rule-based purchasing.	Draft requisitions; detect missing fields.	Route complete requests; resolve eligible exceptions.	Commit approved repeat orders within limits.	Set authority, challenge demand and exceptions.
RFQ & bid comparison	Buyer builds events and comparison sheets.	Draft RFQs; normalize like-for-like bids.	Run eligible events and recommend awards.	Award narrowly eligible cases after enforced checks.	Category strategy, specifications and nonprice trade-offs.
Routine negotiation	Human bargaining; specialist automation exists.	Prepare arguments and term comparisons.	Negotiate specified terms within mandate.	Bounded bilateral agent agreements; conditional forecast.	Reservation policy, strategic terms and disputes.
Supplier risk & SRM	Periodic scorecards and relationship reviews.	Summarize evidence and flag weak signals.	Maintain dossiers; trigger playbooks and follow-up.	Take reversible preapproved mitigations.	Interpret uncertainty; sustain trust and critical relationships.
Strategic sourcing	Category plans and human market judgment.	Research inputs, exposures and should-cost scenarios.	Coordinate analysis and approved sourcing processes.	Selected repeat awards, not unrestricted strategy.	Concentration, innovation, switching and geopolitical judgment.
Demand planning	Forecast systems plus manual overrides.	Explain errors and propose scenarios.	Detect signal changes; run and compare validated models.	Refresh stable segments within service and change limits.	Challenge assumptions, censored demand and structural shifts.

Current-state anchors: Oracle Procurement and sourcing releases, specialist negotiation products, and planning-platform descriptions. [S01, S02, S07, S08, S09, S11, S12, S16, S17] All forward-stage and human-role entries are the report's analytical framework.

IMPLICATION

A single function can contain several maturity levels. The columns describe increasing scope of responsibility, not a requirement to automate every activity.

EVIDENCE-ANCHORED FRAMEWORK

Function transformation matrix • 2 of 2

No stage removes the need for authoritative records, operational confirmation or accountable owners.

Function	Today	AI-assisted	Agentic	Autonomous	Future human role
S&OP / IBP	Periodic consensus and commitment process.	Prepare reconciliation and scenario packs.	Maintain continuous exception and scenario loop.	Execute approved tactical adjustments.	Allocate scarce resources; accept strategic commitments.
Inventory targets	Rules / optimization plus periodic review.	Explain target drivers and data anomalies.	Call optimizers; propose parameter changes.	Replenish eligible stock within constraints.	Service policy, resilience, obsolescence and limits.
Logistics execution	TMS, tracking and manual exception chasing.	Summarize delays; compare approved options.	Coordinate bookings and exception playbooks.	Rebook / reschedule within confirmed rights and budgets.	Disruption command, liability and customer trade-offs.
Global trade / HS	Screening tools and specialist filing work.	Extract and reconcile evidence; propose classification.	Assemble reviewed dossiers; coordinate authorized steps.	Routine permitted filings from validated data only.	Ambiguous classification, origin, value and legal interpretation.
Analytics / KPIs	Dashboards and periodic manual reporting.	Conversational drill-down with source links.	Investigate exceptions; recommend and track actions.	Produce reconciled reports; close eligible action loops.	Definitions, causal judgment and cross-functional decisions.
Network coordination	Portals, EDI / APIs and bilateral planning.	Summarize cross-party events and commitments.	Coordinate tasks across trusted counterparties.	Federated conditional commitments; forecast, not universal.	Network design, disclosure, conflict resolution and authority.
Controls & assurance	Approvals, access controls and audits.	Surface policy conflicts and evidence gaps.	Monitor agent actions and route exceptions.	Enforce deterministic controls; stop unsafe activity.	Accountability, policy changes, independent assurance.

Current-state anchors: inventory, planning, transport, trade, protocol and governance sources. [S02, S03, S04, S05, S06, S09, S15, S19, S20, S21, S22, S25, S27] Autonomous network coordination is a forecast; it is not established by the existence of a messaging standard.

IMPLICATION

The companion landscape matrix provides the same framework at a larger reading size and in CSV format for further analysis.

PROPOSED EVALUATION DESIGN

Acceptance testing and measurement protocol

Test whether the complete workflow deserves execution rights, including its behavior when something goes wrong.

Define the experiment before deployment

Define eligible cases, freeze the policy version and specify the comparison process and intended outcome. Establish representative baselines for completion, error, service, labor and cost. Examine categories, suppliers, languages, values and exception severity separately; a large group of easy transactions should not conceal poor performance on difficult ones.

Where operational and compliance conditions allow, compare randomized eligible cases or well-matched groups. Record other changes in business conditions. A simple before-and-after result can reflect seasonality, a supplier change or process redesign, making it difficult to attribute the difference to the agent itself.

Test family	Examples / pass evidence
Normal case	Correct input, authorized action, expected external acknowledgment and reconciled result.
Ambiguity / missing data	Conflicting quantity, missing specification or stale lead time results in a visible request or escalation.
Authority boundary	Excess exposure, new supplier, prohibited item and split-order attempts are blocked or routed.
Adversarial content	Untrusted document instructions cannot override permissions or cause data disclosure.
Execution fault	Timeout, duplicate event, partial transaction and unavailable tool recover without hidden commitments.
Behavior change	A model, policy or integration update passes regression tests before broader release.

Evaluate the whole outcome

Assess the operational result and compliance with required steps, including abandoned cases, silent failures and human repair. Review a risk-based sample independently. Zero severe incidents in a small pilot cannot establish negligible risk, so combine tests of adverse scenarios, enforceable controls and continued monitoring. Procedure-aware research supports this distinction without supplying a production acceptance threshold. [\[S22\]](#)

Progress from read-only analysis to shadow recommendations, case-approved execution and then unattended action within limits. Each step needs an owner, entry criteria and rollback conditions. Set thresholds using the company's risk appetite and baseline evidence; a borrowed 99% success claim says little about the severity of the remaining failures.

RELEASE CRITERION

Expand execution rights when net outcomes improve at an accepted level of risk, with stable controls and manageable review work. A successful demonstration alone is insufficient.

ANALYSIS + RESEARCH

Choose the decision engine before choosing the agent

Assign language interpretation, prediction, optimization and transaction control to tools suited to those tasks.

Use a language model to interpret incomplete requests, read documents and coordinate tools. Keep constrained allocation and authorization in services that can be tested separately. This division makes it possible to identify whether a failure arose in the numerical plan, the permission check or the interpretation of the request.

Decision type	Supply-chain example	Recommended engine
Stable rules	Eligibility, budget checks, duplicate prevention	Deterministic code and policy engine
Uncertain quantities	Demand, lead times and disruption probabilities	Validated predictive models; calibration
Constrained choices	Capacity, replenishment and network allocation	Optimization and simulation
Unstructured coordination	Clarify requirements; assemble evidence; call tools	Language-model agent with bounded tools
Binding commitments	PO, shipment, payment or declaration state	Authorized system of record + receipt

When a team of agents adds value

Multiple agents can be useful when expertise, access or permissions genuinely differ, as with planning and trade review. Begin with one orchestrator and explicit case state, then test whether specialization improves compliant outcomes after communication, latency, review and support costs. Two agents relying on the same weak evidence do not provide independent assurance.

Cemri and colleagues group failures around system design and specification, inter-agent alignment, and verification or termination. The research examines earlier systems and non-supply-chain tasks. It offers a useful failure taxonomy, but cannot estimate the error rate of live supply-chain agents in 2026. [\[S34\]](#)

Use shared case IDs, versioned inputs, typed handoffs, time budgets, a transaction owner and explicit stopping states. In buyer–supplier exchanges, protect each firm's confidential objectives while maintaining a common record of agreed terms. A protocol can carry the messages without making the parties' commercial interests align. [\[S18, S20, S21\]](#)

PROMOTION TEST

Add another agent only when it improves the tested decision after coordination and review costs are included.

OFFICIAL EVIDENCE + ANALYSIS

Saudi and GCC integration: different rights on different rails

Government procurement, enterprise operations and customs clearance have different decision rights.

Saudi government procurement uses a unified electronic-portal framework. The Ministry of Finance text addresses lifecycle value, fair competition, published criteria and local-content preferences under applicable rules, and prohibits splitting purchases to evade authority limits. Its page links to Etimad. These provisions concern public procurement; they should not automatically be applied to every private purchase. [\[S35\]](#)

A procurement agent must preserve the process

In a public-sector workflow, an agent could check the dossier, trace evaluation to the published criteria and identify absent approvals. It must not invent an exception, alter criteria after bids arrive or divide a purchase to fit its limit. Map the applicable implementing rules with legal advisers and preserve the official submission channel before considering unattended action.

Interface	Credible near-term use	Boundary to preserve
Enterprise ↔ Etimad	Prepare and reconcile authorized procurement evidence	Official criteria, records and approval rights
Enterprise ↔ trade / conformity	Keep goods, certificates and declarations consistent	Different documents carry different authority
Forwarder ↔ terminal / carrier	Reconcile booking, slot, handling and handoff status	An ETA is not capacity or release confirmation
Saudi ↔ other GCC systems	Share permitted events and acknowledged commitments	Separate national permissions and obligations

ATLP describes an Abu Dhabi single window for sea, land, air, industrial and free-zone trade, including port services. Oman maintains its customs and Bayan services. They provide relevant regional infrastructure, while remaining distinct systems rather than evidence of a single autonomous GCC customs operation. [\[S36, S33\]](#)

NIDLDP's Logistics Private Sector Engagement Council describes coordination with the private sector, multimodal integration and logistics challenges in the Vision 2030 context. It supports discussion of how stakeholders cooperate. Its existence does not establish an AI deployment or demonstrate achievement of a future target. [\[S37\]](#)

PRACTICAL REGIONAL PILOT

Pilot one shipment from supplier documents through authorized filing, terminal acknowledgment and receipt. Compare rejected documents, dwell exposure and human repair with the existing process.

DOCUMENTED – NOT TENANT-TESTED

Oracle 26C: capability and execution limits

The useful release detail is what the feature can do, what it needs and where its authority ends.

Area	Documented function	Operational boundary
Inventory planning [S41]	Stockout recommendations, simulation and implementation workflow.	The notes warn that lead-time and upstream supply constraints are omitted. Review and order release remain material.
Transport [S38]	Conversational carrier changes and tender actions.	Permissions and eligible shipment states apply; an action label is not blanket authority.
Warehouse expiry [S40]	Hold affected inbound tasks or cancel at-risk picked/packed outbound units.	Users require the underlying hold/cancel permissions. Quality accountability remains separate.
Trade screening preparation [S39]	Normalize party/contact/address information.	Results are displayed in chat. Direct use in screening is described as a possible future enhancement.

Three implications for vendor evaluation

Test a recommendation that looks attractive but violates a real capacity, quality or lead-time constraint. The receiving planning or execution system should reject it. This establishes whether the workflow can turn a suggestion into a feasible action, rather than relying on how convincingly the suggestion is explained.

For a conversational action, examine its trigger, user or agent identity, approval path, timeout behavior, acknowledgment and reconciliation. These show how much of the workflow is actually unattended. A feature that requires a person to operate it may still be useful, provided that dependence is understood.

Keep future enhancements separate from released functions. Normalized party data displayed in chat may never reach the screening process. An AI feature index does not establish that access, inputs or downstream connections have been configured. Confirm those details for the proposed workflow. [S06, S38, S39, S40, S41]

DEPLOYMENT DECISION

The release evidence supports testing these capabilities. Execution rights should follow results in the actual operating environment, including rejected and interrupted transactions.

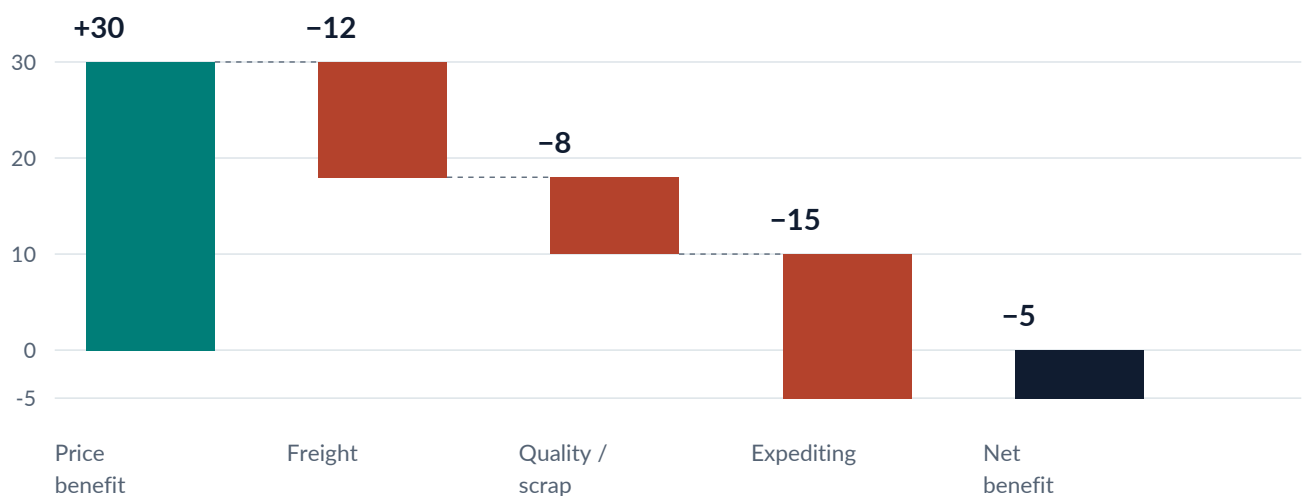
ILLUSTRATIVE – SYNTHETIC

A favorable PPV can conceal a negative decision

A purchasing decision should include the costs it creates after the quoted price is accepted.

Assume 10,000 identical units, baseline unit price SAR 100 and new unit price SAR 97. Specification, currency, quantity and tax basis are unchanged. Under the retained convention, $PPV = (97 - 100) \times 10,000 = -\text{SAR } 30,000$, favorable. All values are synthetic.

Enterprise value bridge • SAR thousands



Calculation: SAR 30,000 price benefit – SAR 12,000 extra freight – SAR 8,000 quality / scrap – SAR 15,000 expediting = **SAR 5,000 net loss** for this purchase. The chart shows value impact; it does not reverse the report's PPV sign convention.

What should the agent do?

Under these assumptions, the award should be rejected despite its favorable PPV. Show alternative suppliers or delivery terms and identify which uncertain cost could change the result. Procurement and finance agree the realized-value calculation; operations then checks whether the proposed service protection was actually achieved.

Authority and quality are independent gates

Quality and authority can rule out a purchase even when its modeled return is positive. Conversely, an authorized emergency response may cost more to protect service. Keep these requirements as hard constraints. For food/QSR decisions, examine cost per usable portion and expiry exposure as well as the quoted price per case.

DECISION RULE

Judge the purchase on enterprise value within its authority, quality and compliance requirements. Confirm the result and required controls; retain material losses and near misses separately so averages do not conceal them. [S22]

PREPRINTS / SIMULATIONS / TEST ENVIRONMENTS

2026 Emerging Supply Chain AI Research

Three controlled studies test different parts of the agent problem; none establishes live-network readiness.

These studies were restored from the 43-reference edition. They examine different tasks and test settings, so their results cannot be combined into one industry accuracy, reliability or savings measure.

Agent bullwhip and reliability

Long et al. • May 2026 • v3 • MIT Beer Game simulation

In the Beer Game simulation, strong average cost results coexisted with ordering instability across runs, facilities and time, including fixed-demand conditions. Repeated sampling did not meaningfully remove the instability. Policy post-training reduced tail events within the study, pointing to a reliability problem that average performance alone would miss. [S49]

Interpretation: Test order variability and severe outcomes alongside mean cost. The study provides simulation evidence, not an observed reliability rate for operating supply networks.

Hybrid agentic supply-chain analytics

Lee et al. • 11 September 2026 • v1 • Replicated multi-echelon inventory analytics

The framework uses a coordinator and specialist agents for exploratory questions and deterministic analytics. In a replicated multi-echelon inventory environment, it achieved competitive accuracy while using fewer input tokens than the single-agent baseline. The result concerns the studied analytical tasks and comparison design. [S50]

Interpretation: Lower token use and competitive test accuracy are useful analytical results. They do not show that a resulting purchase is feasible, authorized or reconciled in production.

Inventory-ordering agents

Yoshizato et al. • 5 February 2026 • v1 • Restricted ordering scenarios / simulation

The ordering design combines structured decision prompts with inventory policy; AIM-RM retrieves similar past experiences. The authors report optimality in a restricted setting and improved results against tested baselines across the studied scenarios. These findings concern the specified ordering problem and simulation conditions. [S51]

Interpretation: Restricted optimality cannot establish robustness under arbitrary conditions. This arXiv full version accompanies an accepted AAMAS 2026 extended abstract; it is not a live ERP deployment.

IMPLICATION FOR THE EXISTING OPERATING MODEL

Compare agents with computational baselines, examine repeated runs and severe failures, and verify the downstream action. Specialization should be judged by the outcome it improves.

PRIMARY SOURCES / EVIDENCE LIMITS

References • 1 of 7

Source IDs remain stable across the report, deck and executive website.

S01 • Oracle — Oracle Fusion Cloud Procurement

[Current product page](#); accessed 18 Sep 2026 | [Product description](#)

Source-to-settle baseline. Not an independently tested capability or a single bundled SCM entitlement.

Open source: www.oracle.com

S02 • Oracle — Oracle Supply Chain Planning

[Current product page](#); accessed 18 Sep 2026 | [Product description](#)

Planning and optimization scope. No universal performance or autonomy inference.

Open source: www.oracle.com

S03 • Oracle — Oracle Inventory Management

[Current product page](#); accessed 18 Sep 2026 | [Product description](#)

Inventory execution baseline; deployment configuration is material.

Open source: www.oracle.com

S04 • Oracle — Oracle Transportation Management

[Current product page](#); accessed 18 Sep 2026 | [Product description](#)

Transportation planning and execution baseline, not a claim that every carrier is integrated.

Open source: www.oracle.com

S05 • Oracle — Oracle Global Trade Management

[Current product page](#); accessed 18 Sep 2026 | [Product description](#)

Trade-compliance capabilities; automated preparation does not transfer clearance authority.

Open source: www.oracle.com

S06 • Oracle — AI for Supply Chain & Manufacturing: Fusion AI feature index

[Release index inspected 18 Sep 2026](#) | [Release documentation](#)

The index includes different quarterly releases. 26D entries are excluded from the current-state assessment.

Open source: docs.oracle.com

S07 • Oracle — AI Agent: Autonomous Sourcing Assistant

[26A release readiness](#) | [Release documentation](#)

Low-dollar, high-volume sourcing; configured policies, jobs, owners and permissions. Not unrestricted procurement autonomy.

Open source: docs.oracle.com

S08 • Oracle — Autonomous Sourcing enhancements

[26B release readiness](#) | [Release documentation](#)

Autonomous sourcing/award attributes can drive approval routing; policy design remains necessary.

Open source: docs.oracle.com

S09 • Oracle — New Fusion Agentic Applications for supply-chain performance

[29 Jun 2026](#) | [Vendor announcement](#)

Announced availability of four additional applications. Availability is not independent proof of operating outcomes.

Open source: www.oracle.com

PRIMARY SOURCES / EVIDENCE LIMITS

References • 2 of 7

Source IDs remain stable across the report, deck and executive website.

S10 • SAP — Supply Chain Management solutions

[Current product page](#); accessed 18 Sep 2026 | [Product description](#)

Broad autonomous-supply-chain positioning; customer metrics are not used as AI causal estimates.

Open source: www.sap.com

S11 • SAP — Joule for supply chain planning

[Current product page](#); accessed 18 Sep 2026 | [Product description](#)

Planning assistance, simulation and agent descriptions. Exact tenant availability was not independently established.

Open source: www.sap.com

S12 • Kinaxis — AI for the supply chain

[Current product page](#); accessed 18 Sep 2026 | [Product description](#)

Maestro Agents and combinations of AI, optimization and simulation. No independent cross-platform benchmark.

Open source: www.kinaxis.com

S13 • Microsoft — Set up supplier communications features of the Procurement Agent

[Updated 26 Aug 2026](#) | [Prerelease documentation](#)

Explicitly labeled production-ready preview; separate agent identity, permissions and optional automatic email sending.

Open source: learn.microsoft.com

S14 • Microsoft — Automate procure-to-pay tasks with Procurement Agent supplier communications

[Updated 27 Aug 2026](#) | [Release-plan description](#)

Supplier communications and purchase-order follow-up scope; not autonomous strategic sourcing.

Open source: learn.microsoft.com

S15 • project44 — AI Agent Orchestration

[Current product page](#); accessed 18 Sep 2026 | [Product description](#)

Exception handling, logistics workflows and procurement-related orchestration. Vendor claims are not audited reliability evidence.

Open source: www.project44.com

S16 • Pactum — Agentic AI for procurement

[Current product page](#); accessed 18 Sep 2026 | [Product description](#)

Policy-driven supplier negotiations. Does not establish universal buyer-agent-to-supplier-agent contracting.

Open source: pactum.com

S17 • Pactum / Honeywell — Honeywell autonomous negotiation case

[Undated case page](#); accessed 18 Sep 2026 | [Vendor-reported deployment](#)

More than 2,500 negotiations reported. No independently audited generalizable savings distribution is supplied.

Open source: pactum.com

PRIMARY SOURCES / EVIDENCE LIMITS

References • 3 of 7

Source IDs remain stable across the report, deck and executive website.

S18 • Anthropic — Building effective agents

19 Dec 2024 | [Primary engineering guidance](#)

Distinguishes workflows from agents and explains complexity, latency and error trade-offs. Not a supply-chain outcome study.

Open source: www.anthropic.com

S19 • NIST — AI RMF: Generative Artificial Intelligence Profile, NIST AI 600-1

26 Jul 2024 | [Voluntary risk-management guidance](#)

Cross-sector lifecycle guidance; not mandatory certification or proof that a particular agent is safe.

Open source: doi.org

S20 • Google Developers — A2A: A new era of agent interoperability

9 Apr 2025 | [Protocol announcement](#)

Agent discovery and task exchange; not legal authority, commercial agreement or guaranteed interoperability in a specific deployment.

Open source: developers.googleblog.com

S21 • Model Context Protocol — What is MCP?

Documentation version 28 Jul 2026 | [Protocol documentation](#)

Interfaces for tools and data; not a transaction-control or regulatory-compliance standard.

Open source: modelcontextprotocol.io

S22 • Cao, Driouich & Thomas — Beyond Task Completion: Revealing Corrupt Success in LLM Agents through Procedure-Aware Evaluation

3 Mar 2026; [arXiv:2603.03116](#) | [Research preprint](#)

Benchmark evidence for procedural evaluation; not a measured error rate for deployed supply-chain agents.

Open source: arxiv.org

S23 • ILO — Generative AI and jobs: A 2025 update

20 May 2025 | [Research brief](#)

Task exposure is not a redundancy forecast, and aggregate exposure is not a supply-chain-specific estimate.

Open source: www.ilo.org

S24 • Brynjolfsson, Li & Raymond — Generative AI at Work

Revised 6 Nov 2024; [arXiv:2304.11771 v2](#) | [Empirical research](#)

5,172 customer-support agents; average 15% issues-resolved-per-hour improvement. Different domain and assisted, not autonomous, work.

Open source: arxiv.org

S25 • GS1 — EPCIS Standard, version 2.0

Ratified Jun 2022; [official reference PDF](#) | [Interoperability standard](#)

Cross-enterprise visibility-event semantics. The PDF inspected identifies itself as version 2.0.

Open source: ref.gs1.org

PRIMARY SOURCES / EVIDENCE LIMITS

References • 4 of 7

Source IDs remain stable across the report, deck and executive website.

S26 • UNCITRAL — Model Law on Automated Contracting

2024 | [Official legal-model overview](#)

Existence of a model law does not establish enactment or its application in Saudi Arabia or another jurisdiction. Final-pass retrieval was blocked; retained only for bounded historical/legal-model context, not a new availability claim.

Open source: uncitral.un.org

S27 • World Customs Organization — What is the Harmonized System?

[Official overview](#); accessed 18 Sep 2026 | [Official nomenclature overview](#)

HS classification rests on a structured nomenclature and interpretation rules; national extensions and current measures require separate verification.

Open source: www.wcoomd.org

S28 • ZATCA — Provide an Issued Declaration

Page updated 9 Sep 2026 | [Government service documentation](#)

Fasah statistical-declaration reuse and AI validation. Not evidence of fully autonomous customs clearance or an AI rollout in 2021.

Open source: zatca.gov.sa

S29 • ZATCA — Provide an Incoming Declaration With an Entitlement Number

Updated 30 Aug 2026 | [Government service documentation](#)

Electronic GCC-related declaration workflow through Fasah; not a single harmonized operational or legal regime.

Open source: zatca.gov.sa

S30 • SASO / SABER — SABER electronic platform

[Official platform](#); accessed 18 Sep 2026 | [Government-linked service documentation](#)

Product and shipment conformity services and Fasah linkage. Future-dated notices are not treated as current requirements. Conformity scope was corroborated on official subpages; homepage retrieval and the earlier Fasah-link statement were not revalidated.

Open source: saber.sa

S31 • SDAIA — Regulation on Personal Data Transfer Outside the Kingdom

[Official regulation portal](#); accessed 18 Sep 2026 | [Official regulation](#)

Personal-data transfers require applicable conditions, safeguards and sometimes risk assessment. No blanket all-data-localization claim.

Open source: dgp.sdaia.gov.sa

S32 • Saudi Vision 2030 / NIDLP — National Industrial Development and Logistics Program Delivery Plan

[Historical 2021–2025 plan](#) | [Government strategy document](#)

Historical industrial, logistics and Fourth Industrial Revolution context. Not a 2026 progress report or a new future target. Final-pass retrieval was blocked; retained only for bounded historical/legal-model context, not a new availability claim.

Open source: www.vision2030.gov.sa

S33 • Oman Customs — Official customs portal and Bayan services

[Official portal](#); accessed 18 Sep 2026 | [Government service portal](#)

Confirms a separate national digital customs platform; no inference of autonomous clearance or shared GCC permissions.

Open source: www.customs.gov.om

PRIMARY SOURCES / EVIDENCE LIMITS

References • 5 of 7

Source IDs remain stable across the report, deck and executive website.

S34 • Cemri et al. — Why Do Multi-Agent LLM Systems Fail?

2025; version 1 | [Research preprint](#)

Failure taxonomy; older systems and non-SCM tasks do not establish a current supply-chain failure rate.

Open source: arxiv.org

S35 • Saudi Ministry of Finance — Government Tenders and Procurement Law: General Provisions

[Official Arabic text](#); inspected 18 Sep 2026 | [Official legal text](#)

Public-procurement scope, portal, criteria and authority. Do not generalize requirements to every private purchase.

Open source: www.mof.gov.sa

S36 • ATLP — Abu Dhabi Advanced Trade and Logistics Platform

[Official service page](#); inspected 18 Sep 2026 | [Official service description](#)

Single-window services and port-related procedures; not evidence of GCC-wide autonomous execution.

Open source: www.atlp.ae

S37 • NIDL — Logistics Private Sector Engagement Council

[Official program page](#); inspected 18 Sep 2026 | [Government program description](#)

Public-private logistics coordination and Vision 2030 objectives; not a measured AI deployment outcome.

Open source: lpsec.nidlp.gov.sa

S38 • Oracle — Shipment Change Assistant

[26C release readiness](#) | [Release documentation](#)

Conversational shipment actions depend on configured permissions and valid shipment states.

Open source: docs.oracle.com

S39 • Oracle — GTM Restricted Party Pre-processing Assistant

[26C release readiness](#) | [Release documentation](#)

Normalized results are displayed in chat; use in screening is described as a possible future enhancement.

Open source: docs.oracle.com

S40 • Oracle — Inventory Expiry Assistant: Enhanced Execution Controls

[26C release readiness](#) | [Release documentation](#)

Cancellation and hold actions require underlying user permissions; not independent autonomous safety certification.

Open source: docs.oracle.com

S41 • Oracle — Inventory Planning Command Center

[26C release readiness](#) | [Release documentation](#)

Recommendations omit lead-time and upstream supply constraints; user review, simulation and order release matter.

Open source: docs.oracle.com

S40 reopening failed in this targeted revision; the prior documented scope is retained without extending its availability claims.

PRIMARY SOURCES / EVIDENCE LIMITS

References • 6 of 7

Source IDs remain stable across the report, deck and executive website.

S42 • OpenAI — Data analysis with ChatGPT

[Official help](#); [inspected 18 Sep 2026](#) | [Product documentation](#)

Code-backed analysis of uploaded files. Review methods and reconcile outputs; file, plan and workspace limits apply.

Open source: help.openai.com

S43 • OpenAI — Deep research in ChatGPT

[Official help](#); [inspected 18 Sep 2026](#) | [Product documentation](#)

Multi-source research with citations. Access depends on plan and permissions; a cited synthesis is not an authoritative commodity forecast.

Open source: help.openai.com

S44 • OpenAI — Enterprise privacy at OpenAI

[Updated 8 Jan 2026](#); [inspected 18 Sep 2026](#) | [Provider data-handling policy](#)

Business data is not used for model training by default. This is not zero retention, permission to upload, or a substitute for company data approval.

Open source: openai.com

S45 • Anthropic — Create and edit files with Claude

[Updated 6 Aug 2026](#); [inspected 18 Sep 2026](#) | [Product documentation](#)

File creation and code-backed analysis; organization controls and network settings matter. Validate formulas, extracted values and permitted access.

Open source: support.claude.com

S46 • Microsoft — Get started with Copilot in Excel

[Official help](#); [inspected 18 Sep 2026](#) | [Product documentation](#)

Editable formulas, tables, charts and workbook analysis. Requires eligible subscription/settings; distinguish workbook editing from ERP authority.

Open source: support.microsoft.com

S47 • Google — Collaborate with Gemini in Google Sheets (Workspace Experiments)

[Official help](#); [inspected 18 Sep 2026](#) | [Product / program documentation](#)

Tables, formulas, analysis and charts. Reviewed page is labeled Workspace Experiments; eligibility, account and language availability vary.

Open source: support.google.com

S48 • Microsoft — Get started with Power Automate approvals

[Official documentation](#); [inspected 18 Sep 2026](#) | [Workflow documentation](#)

Human sign-off orchestration with licensing and environment prerequisites. An approval flow is not by itself an AI decision engine or an ERP transaction right.

Open source: learn.microsoft.com

PRIMARY SOURCES / EVIDENCE LIMITS

References • 7 of 7

Restored studies retain their pinned versions and receive new IDs to avoid collisions.

S49 • Long, Simchi-Levi, Zhu, Su, A. Calmon & F. Calmon — Reliability and Effectiveness of Autonomous AI Agents in Supply Chain Management

16 May 2026; revised 25 May 2026 (v3) | Research preprint / simulation

MIT Beer Game simulation: average performance and decision instability are distinct. Not a live supply-chain ROI or deployment reliability estimate.

Open source: arxiv.org • earlier ID S34 → current S49

S50 • Lee, Inoue, Wang & Gupta — A Hybrid Agentic AI Framework for Intelligent Supply Chain Analytics

11 Sep 2026 (v1) | Research preprint / replicated test environment

Coordinator and specialist agents evaluated in a replicated inventory environment. Reported test accuracy and token efficiency do not establish safe production execution.

Open source: arxiv.org • earlier ID S35 → current S50

S51 • Yoshizato, Shimizu, Higa & Otsuka — AI Agent Systems for Supply Chains: Structured Decision Prompts and Memory Retrieval

5 Feb 2026 (v1) | arXiv full version / simulation; AAMAS extended abstract

Full version of an AAMAS 2026 accepted extended abstract. Restricted ordering scenarios and simulation results are not universal optimality or live implementation proof.

Open source: arxiv.org • earlier ID S36 → current S51

Version reconciliation: The earlier edition contained 43 references. The final register contains 51 unique sources: 40 retained from that edition and 11 subsequent additions. Three earlier contextual items remain outside this narrowly scoped restoration; they are explicitly identified in Reference_Reconciliation.csv. Existing S01–S48 are unchanged.