Deloitte.

The Agentic Supply Chain

Why the Biggest Shift in Supply Chain Operations Is Now Within Reach

Sarthak Pany | Elizabeth Baker | Mahendra Dedasaniya | Nihar Dalmia

Partners, Deloitte Canada | April 2026

Executive Summary

30%

Productivity increase at C.H. Robinson, 2023–2024

The barriers that once prevented supply chain AI from reaching production have resolved. Not partially. Three of four, decisively.

Predictive AI improved the quality of decisions humans made. Generative AI accelerated how humans prepared for them. Agentic AI removes certain decisions from the human queue entirely. Gartner projects that by 2028, one third of enterprise software applications will include agentic AI, enabling 15 percent of day-to-day work decisions to be made autonomously.[20] That is not an incremental capability upgrade. It is a structural shift in how supply chain operations run. Three barriers that prevented organisations from reaching that reality have shifted:

1. Building is no longer the hard part. Low-code agent platforms, including Microsoft Copilot Studio and Google Vertex AI Agent Builder, now allow supply chain professionals to describe a workflow in plain language and deploy a working agent in days, without data science expertise. The builder population has expanded from specialists to the people who understand the operations. The enterprise platforms your organisation already owns, advanced planning systems, ERPs, and best-of-breed supply chain solutions, have embedded agentic AI as generally available capability. The technology question is no longer the primary barrier.

2. The proof that it works now exists at scale. C.H. Robinson's fleet of 30-plus AI agents processed over three million freight shipment tasks in 2025, producing a documented 30 percent productivity increase. Walmart operates autonomous agents across inventory management, supplier negotiation, and logistics, with financial results consistent with AI-driven inventory precision: 5 percent revenue growth on 2.6 percent inventory growth.[17] These are not innovation showcases. They are production deployments with outcomes a CFO can audit.

3. The discipline for capturing value has been proven. Most AI investments failed not because the technology underperformed, but because no one defined the financial outcome before building. The organisations that have scaled AI applied a discipline your leadership team already knows: a specific metric, a defined baseline, a timeline for measurement. Section 7 of this paper codifies this as the Clarity, Intention, and Transparency methodology. It is not new theory. It is capital investment rigour applied to AI deployment.

The fourth barrier has not shifted. More than 70 percent of organisations have deployed AI without redesigning the jobs, workflows, and decision rights it was meant to transform.[6] Individual efficiency gains disappear before they reach the P&L. The organisations that scaled AI made a deliberate choice: they redesigned work around what agents could do and assigned ownership of the outcome to a business leader, not an IT project. That choice cannot be purchased from a vendor. It requires a Chief Supply Chain Officer to make it.

This paper identifies five sequenced decisions that Chief Supply Chain Officers must make to close the gap between pilot and production. The conditions that once blocked those decisions have changed. What remains is the will to make them.

1. The Production Gap

— *Four structural barriers explain why supply chain AI has produced more experiments than enterprise value.*

— *Three have materially shifted. The fourth, redesigning work, remains the domain of leadership, not technology.*

— *Understanding what has changed is the prerequisite for deciding what to do next.*

Barrier 1: The Technology and the Journey Were Not Ready

When most supply chain organisations first experimented with AI, the technology required specialised data science teams, months-long development cycles, and bespoke infrastructure. The journey from proof-of-concept to production was poorly understood, and few organisations had the talent, data maturity, or architectural readiness to navigate it. This barrier has shifted decisively.

The landscape that confronted early adopters was genuinely inhospitable. Predictive AI required hand-crafted feature engineering, custom model pipelines, and continuous retraining, work that demanded scarce data science talent operating in specialised environments disconnected from day-to-day supply chain operations. Deployment cycles measured in months or years were common, and each model served a narrow function: a demand forecast here, a predictive maintenance signal there. The gap between a working model and an integrated operational workflow was vast.

Data challenges compounded the problem. Supply chain data is fragmented across ERP, WMS, TMS, and supplier systems, often in inconsistent formats, with varying latency and questionable accuracy. Building AI on this foundation required extensive data engineering before any model development could begin. Generative AI has begun to ease this burden: large language models can now interpret unstructured data, emails, PDFs, specifications, shipping documents, that previously required manual extraction and standardisation before it could feed analytical models.

The talent constraint was equally real. Traditional AI demanded a combination of deep statistical expertise that few supply chain professionals possessed, and domain knowledge that few data scientists were willing to acquire. This created a persistent bottleneck: business teams understood the problem but could not build the solution; technical teams could build solutions but did not understand the operational context.

What has changed: the technology barrier has largely resolved.

Three shifts have materially reduced this barrier:

1. Enterprise platforms now embed agentic AI. SAP Joule, Oracle Fusion, Kinaxis Maestro, and Blue Yonder ship AI agents as generally available features, though the breadth and maturity of those capabilities vary by platform and use case. Adoption no longer requires a parallel technology stack.

2. Low-code platforms have compressed development cycles. Microsoft Copilot Studio, Google Vertex AI Agent Builder, and Amazon Bedrock Agents enable agent development in weeks, not months, broadening the builder population from data scientists to developer-users. A freight railroad company used Copilot Studio to build an AI assistant that handled over 4,000 customer conversations in its first 45 days.[1]

3. Generative AI has eased the data preparation burden. Large language models can work with imperfect, semi-structured data that would have been unusable in traditional ML pipelines.

However, CSCOs should validate vendor claims carefully: Gartner estimates that only a fraction of self-described agentic AI products deliver truly autonomous capabilities, and predicts over 40 percent of agentic AI projects will be abandoned by 2027 due to costs outpacing value.[46]

Barrier 2: Pilots Were Measured on Activity, Not Financial Outcomes

In our experience, the majority of supply chain AI pilots were evaluated on technical metrics, model accuracy, processing speed, user adoption rates. Rather than on their contribution to the profit and loss statement. This created a structural disconnect: projects could demonstrate impressive technical performance while delivering negligible business value.

40%

of agentic AI projects predicted to be abandoned by 2027 (Gartner)

The root cause was an absence of rigorous value frameworks connecting AI capability to financial outcomes. Organisations treated AI investment as a technology expenditure with uncertain returns rather than as a business case requiring the same discipline applied to capital investment or process redesign. Pilot teams optimised for what they could measure, inference latency, forecast accuracy percentages, automation rates, because no one had defined what financial outcome the pilot was supposed to produce.

What has changed: organisations that have scaled AI treat value as a design constraint, not a post-deployment hope.

The organisations profiled in this paper, C.H. Robinson, Walmart, share a common discipline: they defined the financial outcome before building the solution. C.H. Robinson did not deploy quoting agents to demonstrate generative AI capability; they deployed them to compress quote delivery from hours to an average of two minutes and thirteen seconds, because delayed quotes cost 23–25 percent more in spot pricing.[2] Walmart did not implement self-healing inventory as an innovation showcase; they deployed it to eliminate waste caused by overstock sitting in the wrong stores.[3] The frameworks that enable this rigour, including the Clarity, Intention, and Transparency methodology detailed in Section 7, now exist and have been proven in production.

Barrier 3: Governance and Trust Could Not Keep Pace

As AI tools proliferated faster than organisational policy could accommodate, supply chain functions faced a dual challenge: governance frameworks designed for traditional analytics could not address the risks of autonomous decision-making, and the trust required to delegate consequential decisions to AI had no established process for being earned.

Shadow AI, the adoption of AI tools by individuals and teams outside sanctioned governance structures, emerged as a material risk. IBM's 2025 Cost of a Data Breach Report found that 63 percent of organisations lacked AI governance policies to manage AI and prevent shadow AI proliferation, and that finding predated mainstream agentic deployment.[4] In supply chain, the consequences are operational: an unsanctioned agent modifying replenishment parameters, renegotiating supplier terms, or reclassifying freight without audit trails creates exposure that traditional IT governance was never designed to detect.

Trust is the less visible but equally important dimension. Even where governance structures exist, organisations struggle to define when and how to extend trust to autonomous systems. A procurement agent that performs flawlessly on tail-spend transactions has not earned the trust to negotiate strategic supplier contracts. But without a structured framework for graduating trust, the default is either perpetual human oversight (which eliminates the efficiency case) or premature full autonomy (which creates unmanaged risk).

What has changed: governance now has both regulatory teeth and practical frameworks.

Regulatory frameworks are emerging, led by the EU AI Act (detailed in Section 6), and North American regulation is expected to follow. More importantly, practical governance tools now exist. The Autonomy Ladder introduced in Section 6 provides a structured mechanism for earning and extending trust incrementally, replacing the binary choice between full human oversight and premature autonomy.

It is worth noting that AI governance remains a work in progress. Recent surveys indicate that unsanctioned AI tool usage remains widespread even in organisations with formal policies, though the extent varies significantly by industry and enforcement maturity. Surveys suggest up to 70 percent of enterprises acknowledge that their AI governance is not fully optimised.[31] The shift is not that governance is solved, it is that the tools, frameworks, and regulatory structures to address it now exist. Proactive organisations are building governance capability ahead of enforcement, recognising that the cost of reactive compliance far exceeds the cost of early investment.

Barrier 4: Individual Gains Did Not Become Organisational Gains

This is the barrier that has not shifted. Across industries, a significant majority of organisations, more than 70 percent by recent estimates, have deployed AI without redesigning the jobs, workflows, or decision rights that AI is meant to augment.[6] The result is a pattern that researchers describe as a productivity

70%+

Deployed AI without redesigning the work paradox: individual workers report meaningful time savings, but those savings do not translate into organisational productivity gains because the work itself was never restructured to absorb freed capacity.[7]

The root cause is not data, models, or talent, though leadership often assigns blame there. Pilots stall because no one is accountable for how and when AI should make decisions. Until an organisation explicitly delegates decision rights to AI and changes process at the task level, scaling will stall regardless of the technology's capability.

40%

of AI time savings lost to correcting, verifying, and rewriting outputs

The mechanism is now well documented, and it is specific to organisations that have not redesigned the work. A January 2026 Workday report found that while 85 percent of employees save one to seven hours per week with AI, nearly 40 percent of those time savings are subsequently lost to correcting, verifying, and rewriting low-quality AI outputs, a phenomenon the researchers termed the "AI tax on productivity."[32] A separate study of experienced software developers found that AI tools made them 19 percent *slower* on coding tasks, despite the developers believing they were 20 percent faster, a 39-percentage-point gap between perception and reality.[33] When AI-assisted teams do complete more tasks, the time spent on human approval and review of those tasks can balloon by as much as 91 percent, creating new bottlenecks elsewhere in the workflow.[34]

Consider the difference between layering AI onto a workflow and redesigning the workflow around AI. In traditional tail-spend procurement, a routine purchase follows a predictable path: requisition raised, three-bid sourcing process initiated, buyer evaluates responses, manager approves, purchase order issued. Five handoffs for every transaction, regardless of value. When organisations layer an AI assistant onto this workflow, the assistant may draft the RFP faster or summarise bids more efficiently. But the five-handoff structure remains intact, and the total cycle time compression is marginal.

Now consider the redesigned workflow: an AI agent handles indirect and tail-spend procurement autonomously within policy guardrails, approved supplier list, category spend limits, standard contract terms. Human buyers are redeployed to strategic sourcing and complex negotiations where judgement, relationship management, and commercial creativity create value that AI cannot replicate. For what industry estimates suggest is 60–80 percent of purchase orders that are routine and policy-compliant, the entire approval chain is replaced by governance-at-design rather than governance-at-execution. This is the difference between making existing work faster and making new work possible.

This conclusion is consistent across Deloitte's global research. A separate analysis of agentic AI in manufacturing supply chains reached the same finding independently: organisations should "reimagine and redesign supply chain workflows around the unique strengths of humans and agents rather than simply inserting AI agents into existing operating models."[48]

What has not changed: most organisations have not made the leadership decision to redesign work around AI.

To be fair, there is encouraging momentum on the workforce readiness front. Over 60 percent of corporate VPs and directors attended AI training in 2025, up from 43 percent the prior year, and 73 percent of leaders now actively encourage their teams to use AI.[35] Many organisations have created dedicated AI teams and new roles. But AI readiness is not the same as work redesign. Training employees to use AI tools within unchanged workflows is a necessary step. Not a sufficient one. Fundamental redesign means reengineering processes, decision rights, role structures, and performance metrics around AI capabilities. By that standard, the barrier remains largely intact.

This barrier cannot be resolved by better technology, more data, or improved governance frameworks. It requires a deliberate leadership decision to map existing workflows, identify where human judgement is irreplaceable, define what roles and responsibilities should look like in an AI-augmented operating model, and manage the organisational change required to get there. Section 9 of this paper addresses this directly.

2. From Pilot to Production: Proof Points

— *Two organisations demonstrate what production-grade supply chain AI looks like today.*

— *Both moved beyond pilot by treating AI as an operational investment, not a technology experiment.*

— *Evidence is drawn from publicly available corporate disclosures, verified press releases, and named executive statements. Analysis is Deloitte's own.*

Before examining these proof points in detail, Exhibit 1 maps the breadth of AI use cases now emerging across the four domains of supply chain management, Plan, Source, Make, and Deliver, classified by their level of autonomy: from assistive intelligence through supervised execution to fully autonomous operation. The examples shown are illustrative of capabilities observed across the market; they represent the direction of the field, not an endorsement of specific vendors. Each named use case is referenced to its primary source in the endnotes.

Exhibit 1: AI Use Cases Across the Autonomy Spectrum

From assistive intelligence to autonomous execution across the value chain.[29]

[F] Autonomous (acts within delegated authority) • **[S]** Supervised execution (acts, human oversees) • **[A]** Assistive (recommends, human decides)

PLAN	SOURCE	MAKE	DELIVER
[F] **Autonomous Demand & Inv. Balancing** Ingests real-time data to auto-reforecast SKU demand and trigger stock transfers/POs.	[A] **RFQ Parsing & Entry** Extracts key details from quotes/RFPs to auto-populate procurement fields.	[F] **Intelligent Prod. Rescheduling** Auto-adjusts production schedules during disruptions to minimize downtime.	[S] **Dynamic Route Optimizer** Plans/adjusts transportation routes based on real-time conditions to recalculate optimal paths.
[A] **AI-enabled Demand Planning** Uses historical sales and external/local drivers to generate daily ingredient forecasts.	[A] **PO Issue Resolver** Flags missing items, incorrect quantities, or late shipments in POs and suggests fixes.	[S] **Predictive Maintenance Autopilot** Analyzes IoT data to predict equipment failures, auto-scheduling maintenance and ordering parts.	[F] **Automated Freight Mgmt.** Tracks shipments to automatically initiate exception handling and routine communications.
[F] **Inventory Rebalancing**	[F] **Tail-Spend Negotiation**	[S] **Outside Process Automation**	[F] **Warehouse Ops Scheduling**

Monitors forecasts and real-time inventory to dynamically adjust production and distribution.	Digital buyer for low-value purchases, negotiating via predefined rules and benchmarks.	Automates hand-offs for partially finished goods sent to external vendors, handling documentation.	Forecasts workload and dynamically allocates resources like labor shifts and picking robots.
[F] **Late Supply Mitigation** Detects late shipments to auto-expedite suppliers, re-sequence production, or adjust plans.	[S] **Supplier Discovery** AI-driven global search and vetting of suppliers against complex criteria for sourcing events.	[S] **Real-Time Production Scheduling** Real-time workshop rescheduling considering parts availability, shifts, and sequence constraints.	[S] **Last-Mile Delivery Co-Pilot** Dynamically reorders delivery sequences, consolidates neighbourhood routes, and updates customers.
[F] **Supply Disruption Response** Monitors upstream supply, auto-reassigning materials and booking expedited transport.	[A] **Contract Risk Analyzer** Reviews supplier contracts to flag risks, ensure compliance, and suggest improved clauses.	[S] **Human-Robot Collaboration** Orchestrates fleets of autonomous mobile robots and human workers, coordinating workflows.	[S] **Autonomous Quality Control** Uses generative AI and computer vision to autonomously inspect items for damage/defects before packing.
[S] **Self-Learning S&OP** Runs continuous what-if scenarios on updated data to simulate disruptions and recommend contingency plans.	[F] **Touchless Invoice Processing** Automates matching/ approval of invoices against POs/receipts to flag discrepancies.		[F] **Automated Freight Auditing & Payment** Autonomously audits freight invoices, identifies discrepancies, and processes payments.

Source: Compiled from enterprise announcements, analyst coverage, and verified case studies (2024–2026). Company names are illustrative of use case categories. See endnotes for individual source citations.

C.H. Robinson: Freight Lifecycle Automation

C.H. Robinson, the world's largest freight broker by volume, has deployed a fleet of more than 30 proprietary generative AI agents that automate the complete lifecycle of a freight shipment, from price quoting and load tendering through appointment scheduling, in-transit status checks, and LTL freight classification.[8] The company reports that these agents performed over three million shipment tasks in 2025 alone.[9]

3M+

Shipment tasks performed by AI agents at C.H. Robinson in 2025

The deployment began with the highest-volume, most data-ready workflows. The quoting agent delivers customer-specific spot quotes in an average of two minutes and thirteen seconds, a process that previously required hours of manual work. The company has publicly stated that delays in quoting increase spot costs by 23–25 percent, providing a direct financial rationale for the speed improvement.[10] More than one million quotes have been processed through this agent.

The orders agent converts unstructured emailed tenders into complete, system-ready orders in approximately 90 seconds, processing 5,500 truckload orders daily and saving an estimated 600 hours of manual labour per day.[11] In LTL operations, tandem agents, one contacting carriers to track missed freight, the second reasoning through and executing next steps, run 100 simultaneous carrier interactions, resolving hundreds of daily shipments across more than 11,000 customers. The company reports that 95 percent of manual tracking checks have been automated, with a 42 percent reduction in unnecessary carrier return trips.[12]

C.H. Robinson's Chief Strategy and Innovation Officer has publicly reported a 30 percent productivity increase across 2023–2024, with an additional 15 percent targeted for 2025.[13]

Walmart: End-to-End Agentic Supply Chain

Walmart's approach to supply chain AI is consistent with its broader strategy of building technology as a competitive asset rather than a cost centre. The company has transitioned from model-centric AI, individual algorithms addressing isolated problems, to a system-centric architecture in which purpose-built agents operate across inventory management, demand forecasting, logistics routing, supplier negotiation, and store operations.[14]

The self-healing inventory system automatically identifies overstock in one location and reroutes it to stores with higher demand. This is a fully autonomous workflow: the agent monitors inventory positions continuously, identifies imbalances, and initiates transfers without human intervention for each transaction. Human oversight is reserved for policy exceptions and strategic inventory decisions. Walmart's broader automation strategy, which encompasses robotics, AI-driven software, and autonomous logistics, targets a 20 percent improvement in unit cost averages, with approximately 60 percent of U.S. stores now receiving freight from automated distribution centres.[15]

In procurement, Walmart's deployment of Pactum AI for autonomous supplier negotiation demonstrates what agentic sourcing looks like at enterprise scale. The agents conduct asynchronous, AI-driven negotiations with thousands of mid-tier suppliers simultaneously, achieving a 68 percent agreement rate and 1.5 percent average cost reduction per negotiated contract, with extended payment terms.[16] The programme is expanding to include transportation rate negotiations. These are outcomes that manual procurement processes could not achieve at comparable scale; the autonomous component is not a recommendation engine but a negotiating agent that proposes, counteroffers, and closes within defined commercial parameters.

In 2025, Walmart introduced "Wally," an agentic AI tool for merchants designed to autonomously identify root causes of inventory issues such as out-of-stocks with greater speed than manual analysis allows.[30] The company has also partnered with Gatik to deploy fully driverless middle-mile trucks at commercial scale, the first such deployment in North America, completing tens of thousands of autonomous deliveries since mid-2025.[36]

Walmart's financial results provide corroborating evidence at the enterprise level. In its most recent reporting period, the company achieved 5 percent sales growth with only a 2.6 percent increase in inventory, a ratio consistent with AI-driven inventory precision enabling revenue growth without proportional working capital expansion.[17] Route optimisation AI has saved 30 million unnecessary driving miles, a capability Walmart has since licensed as a SaaS product to other businesses.[18]

Both organisations share a common pattern: they started with the highest-volume, most data-ready workflows, proved value with auditable financial outcomes, and expanded systematically.

3. The AI Stack in Plain English

— *Supply chain AI is not one technology. It is three distinct layers built on a data foundation.*

— *Each layer serves a different purpose. Not every workflow requires all three.*

— *Understanding the stack prevents the two most common mistakes: treating all AI as equivalent, or assuming agentic AI replaces what came before.*

Before examining where AI should live architecturally or how it should be governed, supply chain leaders need a clear taxonomy of what the technology actually does. The market is noisy, vendors label everything "AI" and the distinctions between predictive analytics, generative AI, and agentic AI are often blurred in commercial messaging. This section cuts through that noise.

Exhibit 2: The Supply Chain AI Stack, Three Paths to Value

Each layer is a standalone deployment endpoint. Not every organisation needs all three.

1 Predictive AI	2 Generative AI	3 Agentic AI
Machine learning models that forecast outcomes, detect patterns, and optimise decisions using structured historical data.	Large language models and multimodal systems that synthesise unstructured information, generate content, and augment human decision-making.	Autonomous software agents that perceive, reason, plan, and execute multi-step workflows within defined guardrails, with human oversight at earned autonomy levels.
EXAMPLE Demand sensing models that adjust SKU-level forecasts in real time based on point-of-sale signals, weather, and promotional calendars.	EXAMPLE A planning copilot that summarises exception reports across suppliers, drafts recommended actions, and explains forecast variances in natural language.	EXAMPLE A freight operations agent that monitors shipment exceptions, reroutes loads, renegotiates carrier rates, and updates customer ETAs without human intervention.

Foundation: Decision-Grade Data

All paths require data that is accurate enough for the specific workflow. Not perfect enterprise-wide data quality.

Note: All paths require decision-grade data. Not every workflow requires all three layers, but organisations at enterprise scale will typically deploy across all three as different workflows demand different capabilities.
Source: Deloitte analysis, 2026.

Layer 1: Predictive AI

Predictive AI is where most supply chain organisations started: demand forecasting, lead-time prediction, predictive maintenance, ETA estimation, and risk scoring. These are machine learning models trained on historical data to identify patterns and predict future states. Predictive AI is powerful but passive. It tells a planner what is likely to happen; it does not decide how to respond. For many workflows, this remains the appropriate level of AI capability.

Layer 2: Generative AI

Generative AI operates on a fundamentally different principle. Rather than predicting outcomes from structured data, it generates new content from unstructured inputs: natural language, images, documents, code. In supply chain, its applications include interpreting supplier communications, drafting RFP responses, summarising exception reports, and converting emailed tenders into structured orders. Critically, generative AI also serves as an enabling bridge: it processes the unstructured, messy data (emails, PDFs, contracts, shipping documents) that constitutes a significant portion of supply chain information flows, making the data preparation journey for predictive and agentic AI materially less punishing. Specifically, GenAI can interpret inconsistent supplier records across systems, normalise data formats without manual mapping, and extract structured fields from unstructured documents such as shipping manifests, purchase orders, and quality certificates. This capability is transforming the data readiness equation: organisations that were stalled by fragmented data can now start deploying AI while improving data quality in parallel.

Layer 3: Agentic AI

Agentic AI represents a qualitative shift, not just from insight to execution, but from tools to agents. Three properties distinguish agentic systems from prior workflow automation: persistent goals (agents maintain objectives across sessions and adapt their approach), delegated decision rights (agents act within defined authority without waiting for human instruction), and continuous learning loops (agents improve from each interaction and outcome). Gartner defines agentic AI systems as those that autonomously plan and take actions to meet user-defined goals.[19] Unlike predictive AI (which forecasts) or generative AI (which creates content), agentic AI acts: it executes multi-step workflows, interacts with enterprise systems, makes decisions within defined parameters, and escalates to human oversight when it encounters conditions outside its authority. In supply chain, applications include autonomous supplier negotiation, freight lifecycle management, self-healing inventory, and automated exception resolution.

Gartner predicts that by 2028, one-third of enterprise software applications will include agentic AI, enabling 15 percent of day-to-day work decisions to be made autonomously.[20] A more recent Gartner analysis projects that by 2030, half of all cross-functional supply chain management solutions will use intelligent agents to autonomously execute decisions across the ecosystem.[47] Supply chain providers are already embedding these agents: SAP Joule spans the company's entire cloud portfolio with over 400 AI-driven use cases, Kinaxis Maestro Agents operate natively within the planning platform, and Blue Yonder's Ops Agents continuously monitor and adjust operations.[21]

This shift has a broader implication: agentic AI introduces the concept of a digital workforce, where agents are treated as capacity rather than tools. An agent that processes 5,500 truckload orders daily is not a software feature; it is a workforce resource with measurable throughput, error rates, and utilisation. Organisations that adopt this framing, managing agent capacity alongside human capacity, find it easier to justify investment, measure ROI, and plan for scale.

> ***All three layers require decision-grade data. Not every workflow requires all three, but multi-billion-dollar supply chains will deploy across all three as different decisions demand different capabilities.***

4. Where AI Should Live: Architecture Decisions

— *Architecture is not a technology decision. It is a portfolio decision about where to invest and where to accept what the market provides.*

— *Five lanes. Two tiers. One governance frame.*

— *The most expensive mistake is building what should be bought.*

Deploying agentic AI in supply chain involves a strategic choice that shapes cost structure, speed to value, competitive differentiation, and long-term technical debt. This section introduces a five-lane architecture framework organised into two tiers, Embedded and Build, enclosed by a governance frame that applies regardless of which lane the AI lives in.

Exhibit 3: The Blended Architecture Framework, Where Should AI Live?

Embed for scale. Empower with low-code for speed. Build custom only for crown jewels.

GOVERNANCE & ORCHESTRATION, DECISION RIGHTS | AUDIT TRAILS | SECURITY | INTEROPERABILITY

EMBEDDED AI Capabilities built into existing software

ERP EMBEDDED AI	BEST OF BREED SC PLATFORMS	SPECIALISED AI-NATIVE
e.g., SAP Joule, Oracle Fusion AI AI where transactional data lives. Broad use case coverage across the enterprise cloud portfolio. Scale across existing user base. **CHOOSE WHEN:** The ERP is the system of record and the use case is transactional at scale.	*e.g., Kinaxis Maestro, Blue Yonder, Manhattan Active, Coupa* Deeper domain-specific intelligence for complex SC decisions. Concurrent planning, network optimisation, exception management, procurement. **CHOOSE WHEN:** Domain-specific planning depth exceeds ERP capability.	*e.g., Pactum AI, FourKites, Celonis* Narrow but deep capability for specific high-value problems. Born AI-native; speed to value. Purpose-built for defined workflows. **CHOOSE WHEN:** A narrow, high-value problem has a purpose-built solution.

BUILD Capabilities created by the organisation

COPILOT STUDIO / LOW CODE	CODE FIRST AGENTS
e.g., Microsoft Copilot Studio, Google Vertex AI Agent Builder, Amazon Bedrock Agents Rapid deployment by business users and developer-users. Natural language interfaces ("vibe coding") making agent creation conversational. Deployment in days, not months. **CHOOSE WHEN:** Speed matters, business users can own the solution, and logic is straightforward.	*e.g., Google ADK, LangChain, Azure AI Foundry, CrewAI* Maximum customisation and competitive differentiation. Reserve for defensible strategic advantage. Requires specialised talent. **CHOOSE WHEN:** The workflow is a competitive differentiator and no off-the-shelf solution closes the gap.

Vendor names are illustrative, not endorsements.[43] Most organisations operate across 2–3 lanes.
Source: Deloitte analysis, 2026.

Tier 1: Embedded (Leverage What Platforms Already Provide)

ERP-Embedded AI

For organisations running SAP or Oracle, the first question is whether the use case can be addressed by capabilities the platform vendor is already shipping. SAP's Joule agents now cover production planning, maintenance/operations, and supplier onboarding across key parts of its cloud portfolio.[21] Oracle Fusion SCM offers prebuilt AI capabilities for inventory management and demand sensing.

Choose this lane when the ERP is the system of record and the use case is transactional at scale. Embedded AI capabilities are maturing rapidly but unevenly across vendors. Validate claims against production readiness, not roadmap promises.

Best-of-Breed Supply Chain Platforms

Platforms such as Kinaxis Maestro, Blue Yonder, Manhattan Active, and Coupa offer deeper domain intelligence than ERP-embedded AI for complex supply chain decisions: concurrent planning across multi-tier networks, agent-driven exception management, network optimisation, and autonomous procurement workflows. Kinaxis Maestro Agents handle anomaly detection, prescriptive actions, and supply disruption management natively within the planning platform.[21]

For organisations already operating these platforms, the first step is evaluating the agentic capabilities the vendor has shipped since the last upgrade cycle. Many of these features are already licensed and available.

AI-Native Vendors

A growing category of vendors are built AI-native from the ground up. Pactum AI provides autonomous negotiation,[16] FourKites delivers predictive visibility,[44] and Celonis offers process intelligence.[45] They offer narrow but deep capability for specific high-value problems, with speed to value that broader platforms may not match.

Choose this lane when a narrow, high-value problem has a purpose-built solution and no off-the-shelf option from ERP or best-of-breed vendors closes the gap.

Tier 2: Build (When Standard Capabilities Are Insufficient)

Low-Code / Business-Led Agents

Low-code development platforms (Microsoft Copilot Studio, Google Vertex AI Agent Builder, Amazon Bedrock Agents) enable supply chain teams to build targeted agents in days or weeks rather than months. These platforms are designed for business-led development: supply chain professionals who understand the workflow can build, test, and iterate on agents without deep software engineering expertise.

Increasingly, building these agents is becoming conversational. Tools like Copilot Studio allow supply chain professionals to describe what they need in plain language and iterate in real time, a shift sometimes called "vibe coding" that is making business-led development accessible to a broader population with each platform release.[1] For lightweight use cases, the barrier to entry is approaching zero.

Choose this lane for high-volume, human-heavy workflows where adoption speed matters, business users can own the solution, and the logic is straightforward: customer service triage, shipment visibility queries, exception routing, workflow assistance.

Code-First Agents

Custom agent development using agent development kits, LangChain, Google Gemini ADK, Microsoft Azure AI Foundry, offers maximum flexibility and control. C.H. Robinson’s fleet of 30+ agents was built using this approach.[22]

Choose this lane for the roughly 5 percent of workflows that genuinely differentiate the business. The investment must be justified by the strategic value of the capability: code-first development requires higher upfront cost, longer time to value, and specialised in-house talent. For use cases addressable by embedded or low-code solutions, defaulting to code-first consumes resources that should be allocated elsewhere.

The Governance Frame

Governance is not a lane, it is the frame around all five. Decision rights, audit trails, security, and interoperability requirements apply regardless of which lane the AI lives in.

The governance implication for architecture: policy should establish a governance gate that requires explicit justification before custom development is approved for any use case addressable by embedded or low-code solutions. Without this gate, the natural tendency of development teams is to custom-build, because custom builds are more interesting, more controllable, and more defensible to budget committees than configuration decisions. The gate prevents this default from consuming resources that should be allocated to higher-priority workflows.

The architecture decision is not build versus buy. It is: where does custom investment create defensible advantage, and where does it create unnecessary cost and complexity?

5. Data Foundation: Usable Beats Perfect

— *Data quality is the most cited barrier to AI success in supply chain, and the most misunderstood.*

— *The standard is not perfect data. It is decision-grade data: fit for the specific workflow being automated.*

— *Organisations that wait for enterprise-wide data perfection wait indefinitely.*

The question is not whether data quality matters, but what standard is required to start: decision-grade data, good enough to support a specific delegated decision safely.

The cost of inaction on data readiness is quantifiable. IBM's Institute for Business Value found that unresolved data issues can reduce AI project ROI by 18 to 29 percent.[37] But the cost of waiting for perfect data is equally real: indefinite delay.

Decision-Grade Data: A Practical Standard

The practical alternative to the "fix all data first" approach is a concept we call decision-grade data, good enough data to support a specific delegated decision safely. Decision-grade data is not a compromise but a precision standard: data that is sufficiently accurate, timely, complete, and contextualised for the autonomous workflow it serves, without requiring enterprise-wide perfection.

A replenishment agent needs accurate on-hand inventory positions, reliable lead times, and current demand signals. It does not need perfect supplier financial data, complete product master records, or enterprise-wide cost accounting accuracy. Defining the data boundary per workflow. Rather than per enterprise, is the shift that enables organisations to start.

The operational tool to make this standard actionable is the Minimum Viable Dataset: the smallest collection of high-quality, relevant data needed to automate the first decision loop. For inventory placement, this might be item attributes, dimensions, shelf life, weight, inventory position by location, and near-term demand signals. If those fields are reliable, you can start.

An important caveat: starting with a Minimum Viable Dataset is not permission to neglect the broader data foundation. A quick win must be followed by systematic expansion of data quality and integration, otherwise organisations risk a patchwork of narrow, fragile solutions that work in isolation but cannot scale. The discipline is to start narrow and prove value, while concurrently laying plans to expand the data foundation so that today's pilot becomes tomorrow's enterprise platform, not a one-off deployment.

Use Case First, Data Second

The sequencing mistake most organisations make is treating data readiness as a prerequisite for use-case selection, when the causality runs the other way. The use case defines which data matters. Defining the use case first is not a shortcut; it is the only way to scope remediation work precisely enough to be executable. Organisations that start with a data quality programme in the abstract rarely finish it. Organisations that start with a specific workflow to automate know exactly what done looks like.

Organisations that have scaled AI consistently report that defining the Minimum Viable Dataset forces the right conversations between supply chain, IT, and data teams earlier and more concretely than any enterprise data programme. It converts an open-ended quality problem into a scoped gap-closure exercise with a finish line.

Architectural Approaches to Data Fragmentation

Leading organisations are addressing data fragmentation through pragmatic architectural approaches that do not require multi-year enterprise overhauls. Supply chain data fabrics, unified logical data layers over existing siloed systems, enable AI agents to access data from ERP, WMS, TMS, and external sources without requiring data migration. Kinaxis and Databricks have partnered to build a supply chain data fabric that connects planning engines directly to lakehouse-scale data infrastructure. Blue Yonder and Snowflake offer a comparable integration for logistics and fulfilment data. SAP's Business Data Cloud is pursuing a similar architecture within the SAP ecosystem.[23] These represent a pragmatic recognition that AI deployment cannot wait for system consolidation to be complete.

---

6. Governance: From Managing Models to Delegating Decisions

— *As AI moves from advisory to agentic, governance must shift from model oversight to decision delegation.*

— *The Autonomy Ladder provides a practical framework for earning and extending trust incrementally.*

— *Regulation is no longer theoretical: the EU AI Act becomes fully enforceable in August 2026.*

The governance challenge for agentic supply chain AI is fundamentally different from governing predictive models or generative AI assistants. Predictive models produce forecasts; governance means monitoring accuracy. Generative AI produces content; governance means reviewing for hallucination and bias. Agentic AI executes decisions, and governing agents means defining what decisions they may make, under what conditions, with what escalation paths, and with what audit trails. The most effective governance models are not control functions that slow deployment. They are decision acceleration mechanisms: clear escalation paths, predefined guardrails, and fast exception handling that give organisations the confidence to delegate at speed.

The Autonomy Ladder is a practical framework for managing this transition. Rather than treating autonomy as a binary, either human-controlled or fully autonomous, the Ladder defines three graduated tiers that allow organisations to earn autonomy incrementally, based on demonstrated performance and bounded risk.

> ***Every agent should start at human-in-the-loop. Autonomy is earned by demonstrating reliability within bounded conditions. Not granted as a default. The goal is to earn autonomy quickly for high-performing use cases, rather than granting blanket autonomy overnight.***

Exhibit 4: The Autonomy Ladder, Governance by Decision Type

Start with control. Delegate as trust is established.

Tier	APPLIES WHEN	THRESHOLD	EXAMPLES
HUMAN-IN-THE-LOOP Full Oversight Agent recommends; human approves before any action is taken. *EU AI Act high-risk: mandatory human oversight*	Irreversible or high-stakes decisions with significant financial exposure, safety or regulatory consequence.	Senior approval required No auto-approve threshold	Large purchase orders Supplier contracts Safety-critical rerouting
HUMAN-ON-THE-LOOP Supervised Execution Agent executes; human monitors retrospectively and can intervene.	Partially reversible, moderate financial impact, pattern-based decisions.	Manager review: $25K+ Retrospective audit	Replenishment triggers Appointment scheduling Route optimisation
HUMAN-OUT-OF-THE-LOOP Earned Autonomy Fully autonomous execution; human oversight for policy violations only.	Generally reversible, low financial exposure, constrained scope.	Auto-approve: under $5K Log-only monitoring	Invoice matching Status updates Data enrichment LTL classification

Autonomy is always reversible. Any agent can be returned to a higher-oversight tier at any time.

Source: Deloitte analysis, 2026. Financial thresholds are illustrative. EU AI Act: Regulation (EU) 2024/1689.

Three Tiers of Autonomy

Human-in-the-Loop: Full Oversight

At the highest oversight level, every significant agent action requires explicit human approval before execution. This tier is appropriate for high-value, irreversible, or novel decisions: large purchase orders, strategic supplier contract commitments, safety-critical logistics rerouting, and any decision that creates material financial or reputational exposure.

Human-on-the-Loop: Supervised Autonomy

At the supervised tier, agents execute routine decisions within defined parameters, while humans monitor behaviour retrospectively and intervene when exceptions arise. An agent might autonomously approve purchases up to $5,000 from a vetted supplier, while anything above $25,000 requires a senior manager's approval.

Human-Out-Of-The-Loop: Earned Autonomy

At the highest autonomy level, agents operate with full independence within a tightly defined action space, based on an established track record. Trust is earned through performance, not declared by policy.

Governance in Practice: Regulatory Context

The EU AI Act is the world's first comprehensive AI regulation. Its high-risk classification applies primarily to AI systems used in the management of road traffic and in the supply of critical infrastructure (Annex III). For consumer products and retail supply chains, the exposure is typically indirect but real: your logistics providers and 3PL partners operating in Europe will be subject to these requirements, and your governance frameworks must accommodate their compliance obligations.

The mandatory obligations for high-risk systems include: 1) continuous risk management throughout the AI lifecycle; 2) strict data governance for training data; 3) comprehensive technical documentation; 4) automatic logging for traceability; 5) human oversight mechanisms, including the ability to intervene or shut down the system; and 6) appropriate accuracy, robustness, and cybersecurity standards.[5] Penalties for serious breaches reach 35 million euros or 7 percent of global annual turnover. High-risk requirements become fully enforceable on August 2, 2026.[24]

For Canadian supply chain leaders, the domestic regulatory picture is still forming. Canada's AIDA died on the order paper in January 2025, and no replacement federal AI legislation has been introduced.[42] Provincial frameworks are partially filling the gap (Quebec's Law 25, Ontario's Employment Standards Act amendments), but these are narrow measures, not comprehensive AI governance.[38]

The pattern from data privacy regulation remains instructive. GDPR established the standard; subsequent North American frameworks adopted and adapted its core principles. AI governance is following the same trajectory. Organisations that align their internal AI governance frameworks with the EU AI Act's principles today will be better positioned regardless of which jurisdiction regulates next. The regulation is evolving; the direction is not.

Governance is not a constraint on AI value. It is the mechanism by which value is sustained and trust is earned.

7. Measuring What Matters: Clarity, Intention, and Transparency

— *The gap between AI aspiration and AI value is not a technology gap. It is a measurement gap.*

— *Clarity, Intention, and Transparency provide a structured methodology for converting AI investment into auditable financial outcomes.*

— *Proven value funds the next use case. The framework contains its own flywheel.*

74% / 20%

Aspire to AI revenue growth vs. achieved it

Seventy-four percent of organisations aspire to use AI for revenue growth. Twenty percent have achieved it.[25] The gap is not explained by technology limitations or data quality, it is explained by the absence of a structured approach to defining, designing for, and tracking value from AI investments. Notably, recent industry research found that high-performing AI companies were nearly three times more likely to significantly modify their workflows alongside AI deployment, reinforcing that value realisation requires organisational change, not just technology.[39]

The Clarity, Intention, and Transparency framework addresses this gap directly. It is not a reporting tool or a dashboard, it is a methodology for embedding value discipline into every phase of an AI deployment, from initial ideation through ongoing operations. Critically, the framework contains a built-in feedback mechanism: proven value from one use case funds and justifies the next, creating a self-funding cycle that sustains momentum beyond the initial pilot.

Clarity: Define Value Before Building

Clarity is the discipline of defining, decomposing, and dollarising the value case before any development begins. It answers four sequential questions:

Value Driver, What specific business outcome does this use case target? Not “improve supply chain efficiency” but “reduce working capital tied up in inventory by 15 percent.”

Mechanism, Through what operational levers does the AI create that outcome? For inventory reduction, the levers include safety stock optimisation, cycle stock policy changes, and excess and obsolete identification.

Impact, What is the quantified financial impact of each lever? C.H. Robinson’s quoting agents address a specific mechanism: every delay in quoting increases spot costs by 23–25 percent.

Dollarisation, What is the annual financial value, expressed in a currency the CFO recognises? For a company with $300 million in working capital, a 15 percent reduction releases $45 million, board-level math.

To illustrate how Clarity works in practice, consider Walmart's self-healing inventory capability through the four-question lens. The *value driver* is reducing waste and working capital tied up in misallocated inventory. The *mechanism* is an agent that autonomously identifies overstock in one location and reroutes it to stores with higher demand via inter-store transfers. The *impact* is reduced markdowns, improved availability without proportional inventory growth, and faster correction of imbalances. The *dollarisation* is expressed through Walmart's financial results: sales growth of 5 percent with only 2.6 percent inventory growth, consistent with AI-driven inventory precision releasing working capital at scale.[17]

Intention: Design for Value, Not Features

Intention translates Clarity into execution discipline. It operates in three phases:

Design, Decompose the use case into functional, UX, and technical design, with every component traceable to the value driver.

Plan, Categorise features as fundamental vs. enriching. Prioritise what drives direct value versus what is complex to build.

Scale, Go deep (advanced capabilities), go far (additional users/geographies), go broad (adjacent workflows).

A food manufacturer used this approach for deployment planning and global inventory balancing. The value driver was clear: ensure the right products are in the right distribution centre at the right time. The initial release was scoped to deliver $1 million in annual payback, a focused, falsifiable target.[26]

Transparency: Track, Share, and Sustain

The tracking mechanism must be defined before deployment, not after. The sharing mechanism is equally important. Value that is tracked but not communicated creates no organisational momentum. Weekly or monthly governance reviews that include AI value reporting build the institutional confidence required to approve subsequent investments.

Exhibit 5: The Value Framework, Clarity × Intention × Transparency

Three disciplines that must be applied simultaneously for AI investment to compound.

CLARITY	INTENTION	TRANSPARENCY
Define value before development begins	*Engineer value into the deployment*	*Track value and share results*
Value driver	**Design**	**Track**
What operational metric improves	Workflow integration, role redesign	Continuous measurement at workflow level
Mechanism	**Plan**	**Share**
How AI changes the workflow	Phased rollout, baseline, success criteria	Report to leadership, fund the next case
Expected impact	**Scale**	
Measurable outcome	Expansion triggers, cross-function adoption	
Dollarisation		
Financial value quantified		
FAILURE MODE	FAILURE MODE	FAILURE MODE
Investment without a falsifiable value hypothesis	*Deployment that works technically but never reaches the workflow*	*Value achieved but never measured or communicated*

Proven value funds next use case »»»

Source: Deloitte analysis, 2026.

> ***Organisations that track and prove value from one AI use case create the financial and institutional foundation for the next. This self-funding cycle is the mechanism that separates organisations with one successful pilot from organisations with a scaled AI operating model.***

A critical question the CIT framework forces is: who owns the value? In most organisations, AI initiatives are funded by IT and measured by technology teams, creating a structural disconnect between the people deploying the technology and the people accountable for business outcomes. Leading organisations assign value ownership to the business function, not IT. The supply chain leader who benefits from the automated workflow owns the value case, tracks the metrics, and reports the results. IT enables the deployment; the business owns the outcome.

When Value Materialises

In our experience, the timeline for AI value realisation in supply chain follows a pattern that is more predictable than most organisations expect, provided the Clarity, Intention, and Transparency disciplines are in place from the start.

Low-code agents and embedded platform capabilities typically deliver measurable operational improvements within 4–12 weeks of deployment. Custom-built agents addressing more complex workflows typically require 3–6 months to reach stable production performance. Enterprise-wide operating model transformation is a 12–24 month journey, with value compounding as each successive workflow is automated.

Put simply: how fast you realise value depends on how fast you make the decisions described in Section 9, not on how mature the technology is.

8. The Competitive Divergence

— The gap between supply chain AI leaders and the field is measurable, compounding, and widening.

— Leaders are not doing more AI. They are doing AI differently.

Top-performing supply chain organisations invest in AI at more than twice the rate of low performers, and the differential is widening.[27] A 2025 Deloitte Global CPO survey found that 96 percent of digital leaders achieved or surpassed cost savings targets, versus 80 percent of other organisations.[40] Yet over 80 percent of companies report no measurable productivity impact from AI despite widespread adoption.[41] The gap is explained by *how* AI is deployed, not *whether* it is adopted.

Exhibit 6: Competitive Divergence, Leaders vs. the Field

What separates the 25% reporting transformative AI value from the 75% still experimenting.

LEADERS	THE DIFFERENTIATOR	THE FIELD
Deploy AI into redesigned workflows and role structures	Redesign the work	Deploy AI into unchanged workflows and existing role definitions
Track and prove value at the individual workflow level	Measure at workflow level	Track value at programme or department level, if at all
Build cross-disciplinary teams: supply chain, data, and AI together	Cross-functional integration	Keep AI initiatives within IT or digital functions
Invest in upskilling at scale alongside every deployment	Upskilling investment	Defer upskilling to HR or individual initiative
Commit to production architecture from the outset	Production commitment	Remain in pilot or proof-of-concept without a production path

Every differentiator is a decision. The gap compounds with every planning cycle.

Source: Deloitte analysis based on enterprise deployments, analyst coverage, and client engagements, 2026.

Exhibit 6 captures the five behaviours that consistently distinguish leaders from the field. Every differentiator is a leadership decision, not a technology choice. Leaders redesign work rather than layering AI onto existing processes. They measure value at the workflow level rather than the programme level. They build cross-disciplinary teams that combine supply chain, data, and AI expertise. They invest in upskilling at scale alongside every deployment. And they commit to production architecture from the outset rather than remaining indefinitely in pilot.

This pattern has precedent. As late as 2000, personal computers showed marginal productivity gains that were more than offset by IT spending. The transformation came when companies redesigned business processes around digital capabilities, giving rise to the digital-first companies that dominated the following two decades. Agentic AI is at the same inflection point: the technology works, but the returns only compound when the work itself is redesigned around it.

The window for competitive parity is narrowing in a specific way. Early AI adopters are accumulating proprietary training data, refined decision playbooks, and institutional governance competence that later entrants cannot acquire quickly. A company that has operated autonomous replenishment for eighteen months has not merely automated a workflow: it has built a feedback loop that has trained its agents on its specific network, its specific supplier base, and its specific demand patterns. That institutional knowledge cannot be replicated by buying the same platform a year later. C.H. Robinson illustrates this directly: 30 percent productivity improvement in the first year, with an additional 15 percent targeted in the second, a compounding return that reflects accumulated agent training data and refined operational playbooks, not just continued deployment.[13] The same dynamic applies to governance: organisations that have run graduated autonomy frameworks for multiple use cases have developed the operational reflexes, the escalation protocols, and the leadership trust required to deploy the next agent faster. The compounding effect means the gap between leaders and the field does not close at the same rate that laggards accelerate. It widens. Each planning cycle in which the decision is deferred is a cycle in which the leader extends the advantage.

> ***The window for competitive parity in supply chain AI is narrowing. Agentic AI creates self-reinforcing advantages: proprietary data, refined playbooks, institutional governance competence. This is a board-level strategic priority, not a technology initiative managed within IT.***

9. The CSCO Agenda: Five Decisions That Cannot Wait

— *This section is the paper's sharpest and densest. It is written for the supply chain leader who has read the evidence and is asking: what do I do on Monday morning?*

— *Five decisions, in sequence. Each has a clear owner, a defined scope, and a "done when" test.*

— *These are not aspirational goals. They are concrete, executable commitments.*

The preceding sections have established that the conditions blocking production have materially changed for three of four structural barriers. What remains is the set of decisions that only supply chain leaders can make.

Decision 1: Define the Data Boundary for the First Autonomous Workflow

The first decision is deliberately narrow: select one workflow to automate and define exactly what data it requires. This means defining what decision-grade data means for that specific workflow and establishing the minimum viable dataset: the smallest collection of fields, at the required accuracy and latency, needed to automate the first decision loop. Start with use-case prioritisation. But recognise that agentic AI capabilities and low-code development paradigms may have changed feasibility assumptions since your last assessment.

The selection criteria should emphasise data readiness, decision reversibility, and financial materiality. Not technical ambition. C.H. Robinson started with truckload quoting: high volume, clean data, clear financial impact. Walmart started with inventory rebalancing.

> **Decision 1 is done when:**
> ✓ Use-case prioritisation has been refreshed against current agentic and low-code capabilities.
> ✓ A single workflow is selected with documented justification.
> ✓ The minimum viable dataset is defined: specific fields, accuracy thresholds, latency requirements.
> ✓ The data gap assessment is complete and remediation is scoped and resourced.

Decision 2: Establish Architecture Policy Before Procurement

Before evaluating vendors or approving development projects, establish a clear organisational policy on when to embed, when to build with low-code, and when to invest in code-first development. The policy should include a governance gate: any proposal for custom agent development must demonstrate why embedded and low-code alternatives are insufficient.

> **Decision 2 is done when:**
> ✓ A written architecture policy defines criteria for embedded, low-code, and code-first lanes.
> ✓ A governance gate requires justification before custom development is approved.
> ✓ Procurement evaluations reference the policy. Not ad hoc technical preferences.

Decision 3: Set Autonomy Policy Before Deployment

Before any agent reaches production, define the organisation's policy on decision delegation. The Autonomy Ladder provides the framework. Every agent starts at human-in-the-loop. The policy should specify the criteria for graduating between tiers.

Decision 3 is done when:

✓ A written autonomy policy defines tiers, thresholds, and graduation criteria.

✓ Every agent deployment specifies its initial autonomy tier and escalation paths.

✓ The policy addresses EU AI Act requirements for organisations with European exposure.

Decision 4: Build a Falsifiable Value Case Before Funding

The 74 percent aspiration versus 20 percent achievement gap in AI-driven revenue growth is not a technology failure, it is a value discipline failure.[25] Before approving funding for any AI initiative, require a falsifiable value case: a specific financial outcome, measured against a defined baseline, with a timeline for assessment.

Decision 4 is done when:

✓ Every funded AI initiative has a documented value case with driver, mechanism, impact, and measurement plan.

✓ Baseline metrics are captured before deployment, not after.

✓ Governance reviews include AI value tracking alongside operational KPIs.

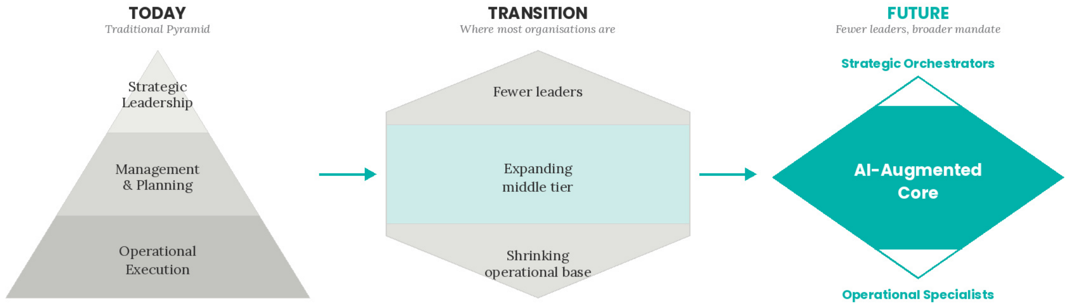
Decision 5: Redesign the Work Before Scaling

This is the decision most organisations have not made. Before scaling any AI deployment, map the roles, responsibilities, decision rights, and performance metrics that will change. The redesign must be specific: which tasks are being replaced, where does freed capacity go, what new responsibilities does the AI create for human workers.

The workforce implications are real. Gartner's research indicates that 51 percent of supply chain leaders expect agentic AI to drive overall workforce reductions, with 55 percent expecting reductions in entry-level positions and 86 percent agreeing that entirely new talent pipeline processes are required.[28]

Exhibit 7: CSCO Role Transformation From Pyramid to Diamond

AI changes what supply chain teams do, and how leadership is structured.



51% of leaders expect overall workforce reductions	**55%** of leaders expect entry-level hiring reductions	**86%** agree new talent pipelines are required

Workforce reductions reflect role transformation and redeployment, not necessarily headcount reduction.
Source: Gartner Future of Supply Chain 2026 (n=509); Deloitte analysis, 2026.

Decision 5 is done when:

✓ Workflow mapping is complete for the first automated workflow.

✓ Freed capacity is explicitly redirected to identified higher-value activities.

✓ Performance metrics are updated to reflect the AI-augmented operating model.

✓ A workforce transition plan addresses upskilling, role evolution, and new capability requirements.

The conditions that once blocked production have changed. What remains are decisions that only leaders can make. An organisation is ready to move from pilot to production when four conditions are met: a governance framework exists that enables fast delegation rather than slow approval; value is owned by the business, not IT, and measured at the workflow level; at least one workflow has been redesigned around AI rather than augmented by it; and leadership can articulate why the competitive cost of delay is no longer acceptable. Looking ahead, the organisations that move fastest will stop thinking in terms of individual use cases and begin designing systems of agents that coordinate across demand, supply, pricing, and risk as an integrated network. That transition is closer than most leaders expect. The decisions in this paper determine whether your organisation will be ready for it.

Continue the Conversation

This paper reflects our current perspective on how AI is reshaping supply chain operations. But the field is evolving rapidly, and the decisions outlined here benefit from dialogue, not just analysis.

If you are navigating the transition from pilot to production and would like to discuss how these frameworks apply to your organisation's specific context, we welcome the conversation.

Sarthak Pany

Partner, Supply Chain & Network Operations

Deloitte Canada | Toronto

spany@deloitte.ca

ca.linkedin.com/in/sarthakpany

Elizabeth Baker

Partner, Supply Chain & Network Operations

Deloitte Canada | Montreal

elizbaker@deloitte.ca

ca.linkedin.com/in/elizabeth-baker-88597b19

Mahendra Dedasaniya

Partner, Supply Chain & Network Operations Leader

Deloitte Canada | Toronto

mdedasaniya@deloitte.ca

ca.linkedin.com/in/mahendra-dedasaniya-79457b32

Nihar Dalmia

Partner, AI & Data

Deloitte Canada | Toronto

nidalmia@deloitte.ca

ca.linkedin.com/in/nihardalmia

Endnotes

[1] CSX use of Microsoft Copilot Studio: Microsoft Copilot Studio customer references, 2024–2025.

[2] C.H. Robinson quoting agent performance and spot pricing impact: C.H. Robinson press release, May 2024; Mark Albrecht, VP AI, C.H. Robinson, 2025.

[3] Walmart self-healing inventory: Walmart Corporate Blog, July 2025.

[4] Cost of a Data Breach Report 2025: 63% of organisations lacked AI governance policies.

[5] EU AI Act: Regulation (EU) 2024/1689. Chapter III (Articles 9–15) details high-risk obligations. Annex III defines high-risk categories including safety components in critical infrastructure. European Commission, digital-strategy.ec.europa.eu.

[6] Proportion of organisations deploying AI without work redesign: Deloitte State of AI in the Enterprise, 7th Edition, 2025–2026.

[7] Supply chain GenAI productivity paradox: Gartner press release, February 5, 2025. n=265 global respondents, August 2024. gartner.com/en/newsroom/press-releases/2025-02-05-gartner-survey-supply-chain-genai.

[8] C.H. Robinson fleet of 30+ AI agents: C.H. Robinson press release, October 31, 2024.

[9] C.H. Robinson 3 million+ shipment tasks: C.H. Robinson press release, April 2025.

[10] Quoting agent performance (2 min 13 sec, 23–25% spot cost impact): C.H. Robinson press release, May 2024.

[11] Orders agent (5,500 daily truckload orders, 600 hours saved): C.H. Robinson press release, October 2024.

[12] LTL exception management (95% automation, 42% carrier trip reduction): C.H. Robinson corporate disclosures, 2024–2025.

[13] C.H. Robinson 30% productivity increase: Arun Rajan, CSIO, April 2025. Microsoft Azure AI Customer Story.

[14] Walmart system-centric AI architecture: Walmart Global Tech Blog, December 2025.

[15] Walmart automation strategy and unit cost targets: Walmart Investor Relations, FY2026 Q4 Earnings, February 2026. 20% unit cost improvement target; 60% of U.S. stores receiving freight from automated distribution centres.

[16] Pactum AI supplier negotiation (68% agreement, 1.5% cost reduction, expanding to transport): Logistics Viewpoints, March 2025; Walmart corporate disclosures, 2025–2026.

[17] Walmart 5% sales growth vs 2.6% inventory growth: Walmart FY2026 financial results, February 2026.

[18] Walmart route optimisation (30M miles saved): Walmart Corporate, 2024.

[19] Gartner definition of agentic AI: Gartner IT Glossary and Forecast, 2025.

[20] Gartner prediction: 1/3 enterprise software with agentic AI by 2028: Gartner Supply Chain Symposium/Xpo 2025.

[21] SAP Joule agents: SAP TechEd, October 2024. Kinaxis Maestro agents: Kinaxis press release, June 2025. Blue Yonder Ops Agents: Blue Yonder ICON, May 2025.

[22] C.H. Robinson technology stack (Azure AI Foundry, LangChain): Microsoft Azure AI Customer Story, 2024–2025.

[23] Kinaxis-Databricks supply chain data fabric partnership: joint press release, March 2025. Blue Yonder-Snowflake logistics data integration: joint press release, January 2025. SAP Business Data Cloud: SAP Sapphire, May 2025.

[24] EU AI Act enforcement timeline and penalties: European Commission, digital-strategy.ec.europa.eu. High-risk obligations enforceable August 2, 2026.

[25] 74% aspire to revenue growth from AI, 20% achieved: Deloitte State of AI in the Enterprise, 7th Edition, 2025–2026.

[26] Food manufacturer $1M payback example: Elizabeth Baker, Deloitte, NISCL presentation, February 2026.

[27] Top supply chain performers invest 2x+ in AI/ML: Gartner, "Supply Chain Executive Report: Future of Supply Chain 2024," press release, February 20, 2024. n=818 practitioners, Aug–Oct 2023. gartner.com/en/newsroom/press-releases/2024-02-20-gartner-top-supply-chain-organizations.

[28] Gartner workforce projections: Gartner Future of Supply Chain 2026, n=509, February 2026.

[29] Exhibit 1 sources: Walmart Q4 FY2026 Earnings Release (Feb 2026); Anaplan PlanIQ; Domino's Pizza UK/Microsoft Dynamics 365 (2023); Coca-Cola/Blue Yonder (2025); Cargill sustainability reports (2025); ServiceNow/Aptiv (2025); Zycus Merlin (2025); SAP/Scoutbee (2025); Coupa platform documentation; Bosch/SAP industry cloud (2025); GE/Siemens MindSphere; Toyota Production System digital transformation (2025); Siemens/Microsoft Industry AI (2025); Blue Yonder platform; FourKites visibility platform; Amazon Robotics (2025); UPS ORION; Turvo/Trimble (2025).

[30] Walmart "Wally" agentic AI for merchants: Walmart Global Tech Blog, 2025; corporate disclosures.

[31] Unsanctioned AI tool usage and governance maturity gaps (70% of enterprises acknowledge governance not fully optimised): Forbes Tech Council, November 2025; Acuvity AI State of AI Security Report, October 2025.

[32] "AI tax on productivity" (40% of time savings lost to rework): Workday AI Productivity Paradox Report, January 2026.

[33] METR developer study (19% slower, 20% perceived faster): Joel Becker, Nate Rush, Elizabeth Barnes, and David Rein, "Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity," METR, July 10, 2025. RCT, 16 developers, 246 tasks. Software engineering study; cited as cross-industry evidence of the productivity paradox. arXiv:2507.09089. metr.org/blog/2025-07-10-early-2025-ai-experienced-os-dev-study.

[34] AI review bottleneck (21% more tasks, 98% more PRs, 91% more review time): Faros AI, "The AI Productivity Paradox Report 2025," June 2025. Telemetry from 10,000+ developers across 1,255 teams. Software engineering study; cited as cross-industry evidence of the productivity paradox. faros.ai/blog/ai-software-engineering.

[35] Corporate AI training momentum (60%+ of leaders attended AI training, 73% encourage teams to use AI): General Assembly, "Corporate Leaders Are Increasingly Upskilling for AI," September 30, 2025. n=651 VPs/directors, US and UK. Coverage: CIO Dive, October 3, 2025. generalassemb.ly/blog/leaders-are-gaining-ground-on-ai-skills.

[36] Gatik autonomous trucks for Walmart: Gatik press release, early 2026. First fully driverless commercial middle-mile operations in North America.

[37] Technical debt and AI ROI: Institute for Business Value, 2025. 86% of executives cite technical debt constraining AI; 18–29% ROI reduction.

[38] Quebec Law 25 (automated decision-making obligations, in force); Ontario Employment Standards Act amendment (AI disclosure in hiring, effective January 2026).

[39] High-performing AI companies are significantly more likely to redesign workflows around AI rather than layer AI onto existing processes: Deloitte, "Bridging the AI Value Gap: Are Team Dynamics the Missing Link?" Deloitte Insights, March 2026. deloitte.com/us/en/insights/topics/talent/ai-roi-and-team-structure.

[40] Deloitte Global CPO Survey 2025: 96% of digital leaders achieved cost savings targets vs. 80% of peers.

[41] NBER study (80%+ of companies report no measurable AI productivity impact): National Bureau of Economic Research, February 2026.

[42] Canada's AIDA (part of Bill C-27) died on the order paper January 2025. No replacement federal AI legislation introduced as of early 2026. Provincial measures (Quebec Law 25, Ontario ESA AI disclosure) provide partial coverage. Parliament of Canada, 2025–2026.

[43] Deloitte provides professional services, including audit and assurance, to certain entities referenced in this publication. References to specific vendors, platforms, or products are for illustrative purposes only and do not constitute an endorsement, recommendation, or preferential treatment. Vendor selection decisions should be based on each organisation's specific requirements and independent evaluation.

[44] FourKites Intelligent Control Tower™ launch and agentic supply chain capabilities: FourKites press release, January 15, 2025. fourkites.com/press/fourkites-introduces-intelligent-control-tower.

[45] Celonis Process Intelligence Platform and AgentC supply chain agent suite: Celonis press release, May 13, 2025. celonis.com/news/press/celonis-showcases-latest-innovations. Platform: celonis.com/solutions/supply-chain-transformation.

[46] Gartner agentic AI project cancellation prediction and vendor maturity assessment: Gartner press release, June 25, 2025. gartner.com/en/newsroom/press-releases/2025-06-25-gartner-predicts-over-40-percent.

[47] Gartner prediction: 50% of cross-functional supply chain management solutions will include agentic AI capabilities by 2030. Gartner press release, May 21, 2025. gartner.com/en/newsroom/press-releases/2025-05-21-gartner-predicts-half-of-supply-chain-management.

[48] Robertson et al., "Resilient by Design: The Agentic Supply Chain," Deloitte Insights, March 2026. deloitte.com/us/en/insights/.../agentic-supply-chain. Henderson et al., "From Vision to Value: A Road Map for Enterprise Transformation in Manufacturing with Agentic AI," Deloitte Insights, September 2025. deloitte.com/us/en/insights/.../agentic-ai-manufacturing.

About Deloitte Canada

At Deloitte, our Purpose is to make an impact that matters. We exist to inspire and help our people, organizations, communities, and countries to thrive by building a better future. Our work underpins a prosperous society where people can find meaning and opportunity. It builds consumer and business confidence, empowers organizations to find imaginative ways of deploying capital, enables fair, trusted, and functioning social and economic institutions, and allows our friends, families, and communities to enjoy the quality of life that comes with a sustainable future. And as the largest 100% Canadian-owned and operated professional services firm in our country, we are proud to work alongside our clients to make a positive impact for all Canadians.

Deloitte provides industry-leading consulting, tax and legal, financial advisory, audit and assurance, and risk advisory services to nearly 90% of the Fortune Global 500™ and thousands of private companies. We bring together world-class capabilities, insights, and services to address clients' most complex business challenges.

Deloitte LLP, an Ontario limited liability partnership, is the Canadian member firm of Deloitte Touche Tohmatsu Limited. Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee, and its network of member firms, each of which is a legally separate and independent entity. Please see www.deloitte.com/about for a detailed description of the legal structure of Deloitte Touche Tohmatsu Limited and its member firms.

To learn more about Deloitte Canada, please connect with us on LinkedIn, X, Instagram, or Facebook.