From a number to a range

Probabilistic forecasting and the future of commodity procurement

How leading CPG organizations are replacing single-point forecasts with probabilistic, explainable decision intelligence, and what it means for sourcing, inventory, and negotiation.

WHITEPAPER

By Aarush Upadhyay
Associate Consultant,
UST

ust.com

UST

Executive summary

Commodity volatility is not a temporary condition procurement teams can plan around and wait out; it is the operating environment. The World Bank's April 2026 outlook projected overall commodity prices to rise 16% in 2026, and disruption to a single trade corridor pushed urea prices from roughly US$400 to more than US$850 per tonne within months.

Yet many procurement organizations still price this volatility with the same tools they used a decade ago: a single-point forecast, refreshed monthly, sitting in a spreadsheet. This white paper argues that the real gap in procurement is not a shortage of data. It is that the data does not connect, and that a single forecast number hides more risk than it reveals. Drawing on Gartner's 2025 research into procurement technology adoption and an anonymised case study from one of the world's largest CPG companies, it sets out why the shift from a single forecast to a range of plausible outcomes is becoming the defining difference between procurement teams that hold a genuine information advantage and those still negotiating from a position of structured guesswork.

The paper closes with a practical view of what comes next: agentic AI systems that don't just forecast a range of outcomes but monitor it continuously and flag the moment a decision is needed, with human validation kept firmly in the loop as a matter of governance, not limitation.

Table of contents

1. Introduction: Volatility is the baseline, not the exception

Commodity volatility remains a live procurement issue, not a periodic disruption. In April 2026, the World Bank forecast overall commodity prices to rise 16% in 2026, driven by a 24% increase in energy prices and a 31% increase in fertiliser prices[1]. The FAO's June 2026 Food Outlook reported that market conditions remained exposed to weather, volatile energy and fertiliser markets, geopolitical tensions, and trade-policy uncertainty[2]. The impact can travel through a supply chain faster than most planning cycles can react: WTO analysis found that disruption to Persian Gulf fertiliser trade pushed urea from around US$400 per tonne to more than US$850 per tonne in April 2026[3]. For CPG procurement teams, input-cost assumptions can shift faster than a monthly spreadsheet, or a negotiation cycle, can keep up.

And yet, even against this backdrop, many procurement teams still make high-value sourcing decisions using single-point forecasts. Mean absolute percentage error (MAPE) measures the average size of forecast errors as a percentage of actual values; it does not measure the share of forecasts that are wrong[4]. The practical problem isn't a single, universal industry error rate. It's that a single forecast number can hide the uncertainty behind it entirely, leaving buyers with no view of the range of plausible outcomes they're actually pricing against.

> **A single forecast number doesn't remove uncertainty. It just hides where it's sitting.**

2. Why data isn't the problem

The real problem in procurement isn't a lack of data; it's that the data doesn't connect. A governed data layer needs to bring together purchase histories, supplier contracts, commodity and energy indices, freight and logistics costs, currency movements, and external signals like geopolitical and event data, and land them in one governed store. Every source should run its own pipeline, so that a problem in one feed can't spread to the others, and every record should carry a note of where it came from and when it arrived.

Gartner's supply chain hype cycle makes the underlying point bluntly: the shift is not toward more forecasting but toward managing uncertainty[5]. In practice, a buyer pricing for the next quarter is often working from monthly-updated batch data, filling the rest with a year of institutional memory. This is more structured guesswork than a forecasting process.

3. The cost of traditional forecasting

Information asymmetry makes the problem worse. Commodity suppliers are sophisticated market participants. They run real-time feeds, staff professional hedging desks, and monitor markets for weeks before a negotiation is even scheduled. By the time a procurement team sits across the table, the supplier already holds the upper hand.

Probabilistic modelling and scenario simulation are standard tools for any commodity trader or asset manager. The industry has widely accepted the resulting information gap as simply the cost of doing business. It shouldn't be.

4. What Gartner is really saying

Gartner's Hype Cycle for Procurement and Sourcing Solutions tracks where the industry actually stands: between 20% and 50% of tracked companies have already purchased or implemented some form of predictive analytics for procurement[8]. But buying the tool and knowing how to use it are different. In practice, buyers often go back to the methods they're used to, because the output isn't trusted, the model isn't understood, or the underlying data was never cleaned well enough to be dependable. The tool gets marked as "implemented." The decisions get made the same way they always were.

In 2024, 71% of CPG leaders reported using AI in at least one business function, up from 42% in 2023[6]. But adoption alone isn't the same as value. The advantage shows up only when organizations embed AI into day-to-day decisions and give procurement teams evidence they can challenge and us, not a black box they're asked to trust.

5. The shift: From a number to a range

The organizations overcoming these hurdles are not the ones with the biggest procurement teams or the most experienced buyers. They're the ones that recognized the future isn't a single, solid number — it's a range of possibilities.

Once a range comes into play, everything downstream of it changes: negotiations, hedging, and planning all shift, along with how much stock a business actually needs to hold. When buyers have a clearer view of the likely price path and the odds of a sharp, random spike, they need less safety stock as protection. That reduces the storage, insurance, handling, and financing costs captured in Carrying Cost of Inventory. Holding less excess stock also means inventory turns over faster, lowering Days Sales of Inventory (DSI). At the same time, the Inventory to Sales Ratio (ISR) should improve because the business is holding a stock level more closely matched to actual demand, rather than tying up cash in inventory that may sit unused.

Compare AI-enabled forecasting against the existing method using the same target, horizon, and error metric. The key shift isn't a headline accuracy figure. It's giving buyers a distribution of plausible outcomes, the drivers behind them, and a clear basis for action. The chapter that follows explains how we applied that approach in an anonymised UST case study from one of the world's largest CPG companies.

6. Case study: Probabilistic forecasting inside a global CPG leader

The client managed billions of dollars in annual raw-material spend, but its procurement teams relied on inconsistent manual forecasts that were often out of date by the time they reached decision-makers. UST set out to close that information gap without requiring a data scientist to sit beside every buyer.

A LAYERED, TRACEABLE ARCHITECTURE

UST built the platform in layers, each doing one job and passing a clean, traceable result to the next. That matters because buyers need to understand where a forecast came from before they can use it with confidence. Instead of presenting a single year-end commodity price, the platform produces a range of plausible outcomes - a likely path, an adverse spike, a softer landing, and the probability attached to each scenario.

That gives the client a figure it can budget against without pretending the uncertainty isn't there.

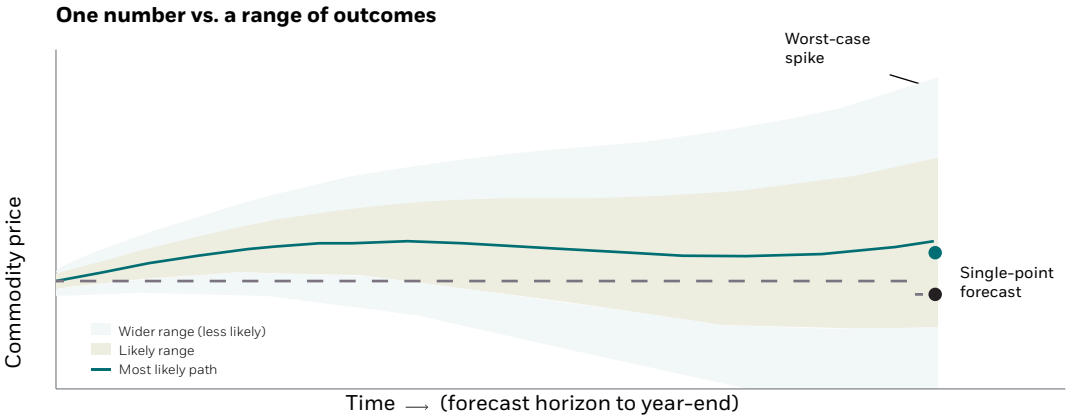


Figure: Illustrative output — a probabilistic price range with scenario bands, replacing a single point forecast (BM)

Underneath, the platform runs in two stages. The first is a deep-learning model that reads the market and the forces that move it (energy, freight, regional supply and demand) along with the other inputs that feed into what a tonne of a commodity ends up costing. This is where the platform works differently from most forecasting tools: many blend every input into a single timeline and forecast from that combined signal. The problem is that blending throws away the detail a buyer actually needs. This platform keeps each driver separate and models how they affect each other, rather than assuming they move independently. By modelling those links directly, it picks up knock-on effects that single-signal models miss, a significant part of why it holds an accuracy advantage over traditional methods.

Each commodity and market is handled on its own terms. Before a model is considered for use, it's trained on historical data and tested against a separate historical period whose outcomes were withheld from training. Models that meet the agreed validation criteria then pass through scenario engines that produce a range of plausible outcomes and their probabilities, rather than a single point estimate. The output is then run through several probabilistic scenario engines, each stress-testing the forecast under different conditions, so what reaches the buyer is a probabilistic range rather than a single number they're simply asked to trust.

EXPLAINABILITY: WHY BUYERS ACTUALLY TRUST IT

Buyers will not use a forecast they cannot explain. The platform was therefore designed to show which drivers moved a forecast, and by how much. The historical evaluation available for this paper covered year-end commodity-price forecasts and data through Q4 2023, reporting an average forecast error of around 6% across the in-scope test set. Because the case material does not disclose the full commodity list, sample size, or tolerance definition, read this figure as a historical back-test result, not live production performance or a universal accuracy claim.

The forecast range communicates uncertainty, while driver-level explanations show how factors such as energy, freight, and currency changed the result. That combination makes the output easier for buyers to challenge, explain, and use directly in a negotiation.

Forecast accuracy is only one part of the case. The available material does not report measured post-deployment changes in downstream procurement, inventory, or service KPIs, so this paper does not claim them as realised benefits. In a full production deployment, the model should also be measured against sourcing-cost variance, inventory carrying cost, Days Sales of Inventory, freight cost per tonne, and On-Time In-Full (OTIF), using clear baselines and consistent attribution. Cross-industry comparisons should only be drawn when the forecast target, horizon, and error metric are genuinely the same.

BUILT TO LAST: AUTOMATION AND SCALABILITY

The platform was designed to refresh as new data arrives, retrain and back-test candidate models, and require human approval before deploying any new model. An ongoing monitoring layer checks for data or model drift and compares the production model against challenger models in real time.

Because the architecture is configuration-driven, the same core platform can extend to additional commodities by changing its inputs, drivers, and validation rules, rather than rebuilding the solution from scratch. That supports scale while keeping each commodity's behaviour and governance requirements distinct, an approach consistent with how UST approaches supply chain visibility and orchestration more broadly across client engagements.

Freight cost per tonne shipped is a directly relevant KPI, since freight is one of the drivers the model tracks explicitly. Better anticipation of price and logistics movement can reduce last-minute surprises and expensive shipments. OTIF should also be monitored, though forecasting is only one of several levers that influence it, so its impact there is indirect rather than direct.

The scale of movement between commodities, and how disconnected their drivers can be, is easy to underestimate. In a single window in 2021[7], palm oil climbed from $0.72 to $1.01 a kilogram, a jump of more than 40%, while crude oil moved the other way, sliding from $68.68 to $60.40 a barrel, and wheat pushed up from $240.29 to $303.01 a metric tonne. Three commodities, the same broad economy, the same window of time, moving in different directions for different reasons. The drivers that move one may not move another, and they don't necessarily stay constant year to year. That volatility, and its unpredictability from commodity to commodity, is precisely why a platform built this way was necessary.

EXTERNAL INTELLIGENCE AND HONEST LIMITS

The platform also connects to a wider intelligence layer that enriches forecasts with external signals such as geopolitical risk and global event data.

Same economy, same window — three commodities, three directions

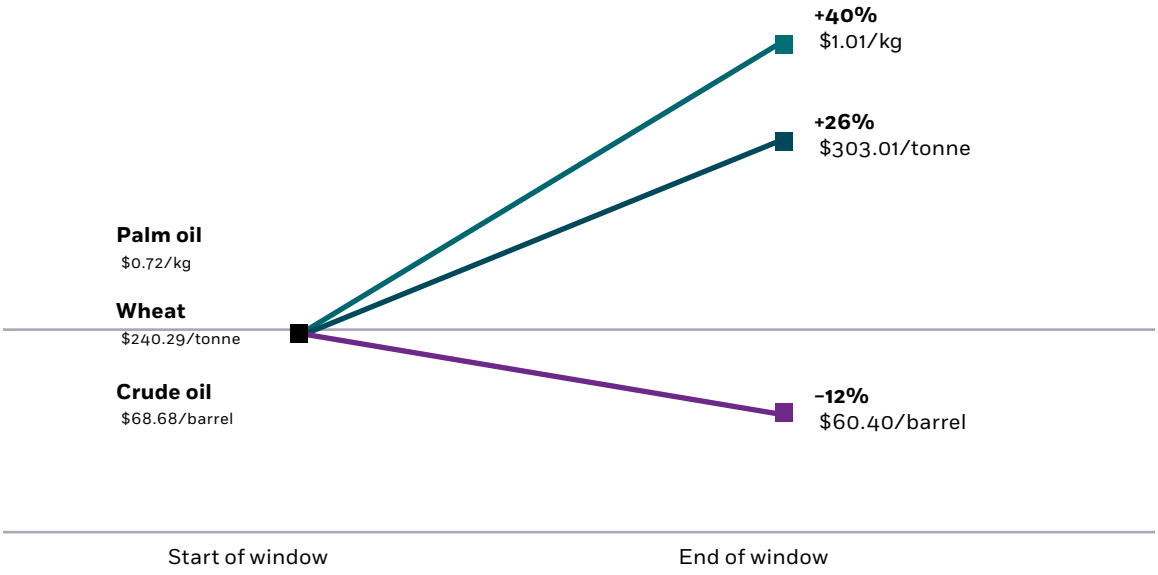


Figure: Illustrative view — external signal enrichment feeding into the probabilistic forecasting layer

A limitation highlighted in the 2025 Hype Cycle for Procurement and Sourcing Solutions by Gartner[8] is worth taking seriously. AI is good at spotting patterns, but it cannot predict a genuinely unseen event, it works from historical data. That is precisely why UST's approach hands back a range with probabilities attached, rather than a single number. A buyer isn't told what will happen. They're told what might happen, and how likely each scenario is, which makes planning considerably more efficient.

Gartner's two 2025 cycles[9] tell the same story from different angles[10]. The supply chain report describes the industry shifting away from point forecasting and toward managing uncertainty, with probabilistic, range-based planning as the way forward. The procurement report goes further on timing, arguing that predictive and prescriptive analytics are no longer bets. They are proven capabilities that should be deployed as soon as possible.

AGENTIC AI AS THE NEXT STEP

- Agentic AI has just entered Gartner's Hype Cycle with under 1% adoption, but it carries a notably high transformational-benefit rating and an expected path to mainstream adoption within the next two to five years. These are systems that act once a forecast has been formed. They watch the market closely, flag when it's time to move, and start drafting the next step before a human even asks.
- The same agent can also track broader operational KPIs. A price spike could trigger an automatic review of purchase orders, inventory coverage, and freight cost per tonne shipped, while a potential supplier disruption could prompt a check against on-time delivery and OTIF exposure, before the system recommends what to do next.
- The same procurement report is unusually direct about where the hype outran the reality[11]. Agentic AI for procurement has fallen into the trough of disillusionment: plenty of firms invested and got mixed results. The reason was simple. Pointing a generic AI tool at messy, contradictory procurement data was never going to work well in the first place. Buying a tool is not the same as gaining an advantage. Between 20% and 50% of procurement teams have purchased predictive analytics[12], which on its own tells you very little, since the purchase and the payoff can be years apart, and in some cases, never arrive at all.

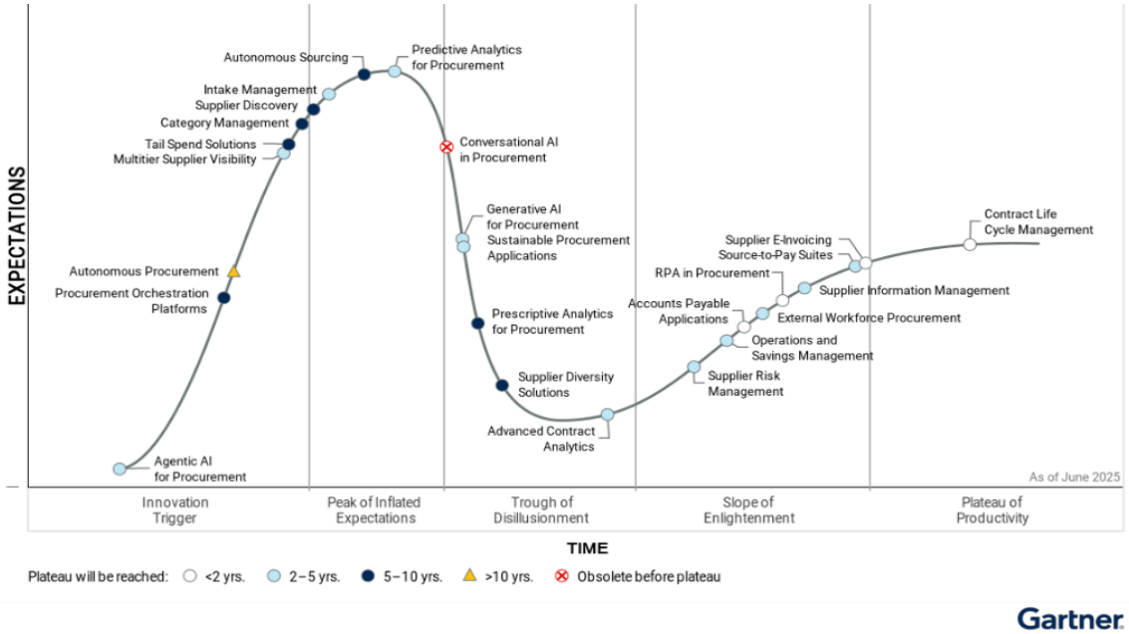
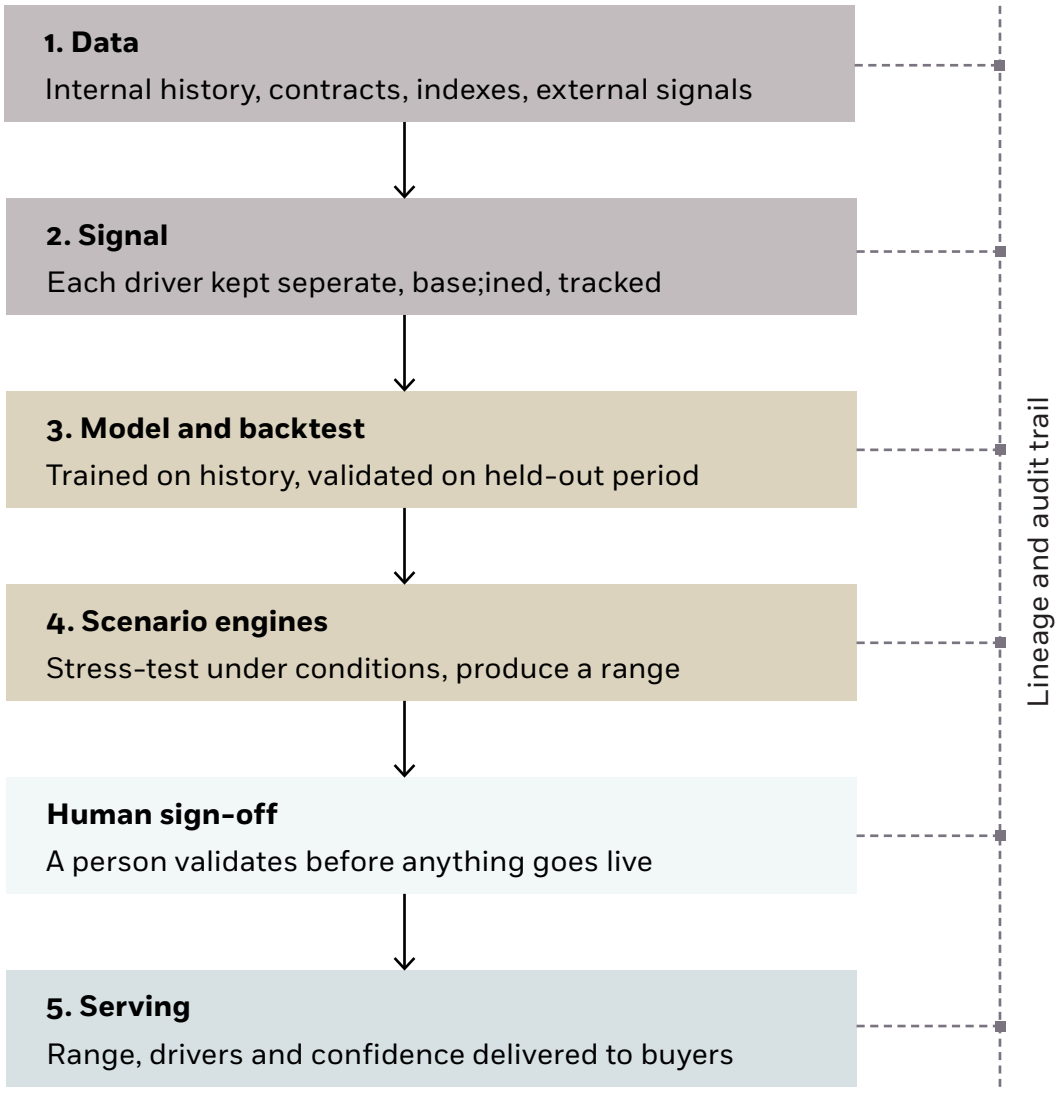


Figure: Illustrative Gartner Hype Cycle positioning for agentic AI in procurement

GOVERNANCE: BUILT IN, NOT BOLTED ON

Governance is built into the platform, not added afterward. No forecast reaches a buyer without a person validating it first, which keeps accountability with the team rather than the model. Every forecast is auditable, tied to the exact data and model version that produced it so that any decision can be explained and defended afterward. A monitoring layer watches both incoming data and the live models, and raises a flag whenever accuracy starts slipping, or a challenger model starts outperforming the one currently in production. And because the platform runs inside the client's own cloud environment, sensitive information (supplier terms, internal financials) never leaves that environment. This combination lets the platform run with real autonomy: every step stays visible, reversible, and owned.



Every layer is traceable to the forecast it produced

Nothing reaches production without validation

Agentic AI is a logical extension of the forecasting platform, not a separate initiative. Once the system can produce a range of plausible commodity-price outcomes, an agent can monitor those ranges, flag material changes, and draft the next action for a buyer to review. Human validation isn't a limitation here; it's the control that keeps accountability with the procurement team. Teams get the most value from AI when they engage critically, challenge its recommendations, and retain decision rights rather than deferring blindly. Zero100[14] cites research showing that manager role-modeling was associated with a 22-percentage-point increase in critical thinking about AI among their team, reinforcing the value of informed, active human oversight, not passive sign-off.

Conclusion: The information gap no longer has to be normal

For a long time, the information gap between commodity suppliers and the procurement teams buying from them was treated as normal, an accepted cost of doing business. It no longer has to be. The competitive edge in procurement is moving away from forecast accuracy as a headline number, and toward a widening divide between organizations that have adopted genuine decision intelligence and those still relying on static, single-number forecasts.

A forecast's accuracy doesn't matter if it stays static inside a report. What matters is whether an improved forecast leads to faster decisions, and whether those faster decisions drive measurable movement across the KPIs that actually run the business: sourcing cost variance, inventory carrying cost, freight cost per tonne, and service levels. That is the real measure of success, and it's the standard UST's procurement and supply chain teams hold this kind of work to.

READY TO MOVE FROM A NUMBER TO A RANGE?

Let's talk about probabilistic forecasting for your procurement organization

If your procurement team is still pricing volatility with a single-point forecast, the gap between you and a genuine decision-intelligence advantage is closer than it looks. UST's supply chain and AI teams work with global CPG, retail, and manufacturing organizations to build explainable, auditable forecasting platforms that hold up in real negotiations, not just a slide deck.

Talk to UST's supply chain experts →

References

[1] World Bank (2026) Commodity Markets Outlook, April 2026. World Bank, 28 April 2026.

[2] Food and Agriculture Organization of the United Nations (FAO) (2026) Food Outlook: Global Food Commodity Market Trends Face Rising Geopolitical and Weather Risks. FAO, 18 June 2026.

[3] World Trade Organization (WTO) (2026) Fertilizer Trade Impacted by Strait of Hormuz Conflict. WTO Secretariat, 10 July 2026.

[4] Hyndman, R.J. and Athanasopoulos, G. (2021) Forecasting: Principles and Practice, 3rd ed., Section 5.8: Evaluating Point Forecast Accuracy. OTexts.

[5] Gartner (2025) Use Range-Based Planning to Manage Supply Chain Uncertainty. Gartner.

[6] McKinsey & Company (2024) Fortune or Fiction: The Real Value of a Digital and AI Transformation in CPG. McKinsey & Company, 3 October 2024.

[7] United States International Trade Commission (USITC) (2022) The 2021 Commodity Price Surge: Causes and Impacts on Trade. USITC.

[8] Gartner (2025) Hype Cycle for Procurement and Sourcing Solutions, 2025.

[9] Gartner (2025) Use Range-Based Planning to Manage Supply Chain Uncertainty.

[10] Gartner (2025) Hype Cycle for Procurement and Sourcing Solutions, 2025.

[11] Gartner (2025) Gartner Says Generative AI for Procurement Has Entered the Trough of Disillusionment, 30 July 2025.

[12] Gartner (2025) Hype Cycle for Procurement and Sourcing Solutions, 2025, 30 June 2025.

[13] Gartner (2025) Gartner Says Generative AI for Procurement Has Entered the Trough of Disillusionment, 30 July 2025.

[14] Zero100 (2026) AI Can Make More Decisions. That Doesn't Mean It Should, June 2026.

About the author

Aarush Upadhyay
Associate Consultant, UST, GB

Aarush Upadhyay is an Associate Consultant at UST, working with clients across procurement, supply chain, and data-driven transformation. His work focuses on turning fragmented operational data into decision-ready intelligence that procurement and supply chain teams can trust and act on. To learn more about UST's approach to AI-powered procurement and supply chain transformation, visit ust.com

Together, we build for boundless impact

Since 1999, UST has worked side by side with the world's best companies to make a powerful impact through transformation. Powered by technology, inspired by people, and led by our purpose, we partner with our clients from design to operation. Our digital solutions, proprietary platforms, engineering, R&D, products, and innovation ecosystem turn core challenges into impactful, disruptive solutions. With deep industry knowledge and a future-ready mindset, we infuse expertise, innovation, and agility into our clients' organizations—delivering measurable value and positive lasting change for them, their customers, and communities around the world. Together, with 30,000+ employees in 35+ countries, we build for boundless impact—touching billions of lives in the process.

ust.com

Version 0101-20260916

U ▪
S T