
RESEARCH ARTICLE

Milestone-Based Architecture for Traceable On-Time-In-Full Delivery Intelligence in Supply Chains

Venu Gopal Kakarla

Independent Researcher, USA

Corresponding Author: Venu Gopal Kakarla, **E-mail:** venugopalkakarla555@gmail.com

ABSTRACT

On Time In Full delivery is one of the most closely tracked measures in global supply chains, but most organizations can only say whether an order missed its target, not where in the fulfillment lifecycle it went wrong. This article proposes a milestone-based architecture that treats fulfillment performance as a traceable lifecycle instead of a single terminal outcome. The lifecycle breaks into measurable checkpoints spanning order acceptance, inventory allocation, picking, dispatch, carrier handoff, delivery confirmation, and proof-of-delivery capture, each carrying a standardized timing and ownership expectation. Drawing on control-tower and performance-measurement research, the article argues that trustworthy On Time In Full reporting depends far more on governed data integration across planning, warehouse, and logistics systems than on any single predictive algorithm. A regulated, multi-region pharmaceutical and medical device supply chain illustrates the design choices throughout. The discussion turns to root-cause analysis, logistics partner collaboration, and the governance tradeoffs that decide whether milestone data becomes something teams can act on or just another dashboard nobody trusts.

KEYWORDS

Fulfillment performance, milestone tracking, root-cause analysis, delivery intelligence, supply chain control tower

ARTICLE INFORMATION

ACCEPTED: 12 September 2026

PUBLISHED: 12 October 2026

DOI: 10.32996/jcsts.2026.8.10.9

1.0 Introduction

Almost every supply chain organization reports On Time In Full (OTIF) delivery. Almost none of them can explain a miss quickly. A shipment either arrives on schedule with the right quantity or it does not, and that binary outcome gets folded into a scorecard that executives review on a cadence far removed from the operations that produced it. When a delivery misses its target, the organization usually knows the fact of the miss well before it knows why. Teams then start a manual investigation across order management, warehouse, and transportation systems that were never built to be queried together. By the time a root cause surfaces, next month's shipments are often already repeating the same pattern.

This article introduces a milestone-based architecture for On Time In Full delivery intelligence. Instead of treating delivery performance as one downstream number, the architecture breaks the fulfillment lifecycle into discrete, measurable checkpoints, each carrying an expected timing, an actual timing, a responsible party, and, where relevant, an exception reason. This is not just a more detailed dashboard. It is a structural shift away from outcome reporting and toward traceable fulfillment intelligence, where the organization can point to the specific stage where performance broke down instead of just the fact that it broke down somewhere.

What follows positions this approach against existing control-tower and performance-measurement research, walks through the milestone-based architecture and the data integration it requires, treats standardized expectation modeling as a governance responsibility rather than a configuration detail, and closes with root-cause analysis, partner collaboration, and the real limits of the approach as described here. A global pharmaceutical and medical device supply chain grounds the discussion throughout, chosen because regulated, multi-region fulfillment operations expose data fragmentation in its most acute form.

2.0 Literature review

2.1 *Fragmented fulfillment data*

Supply chain performance measurement has moved well beyond simple cost and volume tracking over the past two decades. The metrics literature has repeatedly made a related point: measurement systems fail when they are not grounded in an integrated view of the processes they claim to measure (Gunasekaran et al., 2001). Gunasekaran, Patel, and Tirtiroglu's early framework for supply chain performance measures distinguished strategic, tactical, and operational metrics and argued that operational-level metrics are the most vulnerable to poor data integration, since they depend on granular, cross-functional event data rather than periodic aggregates (Gunasekaran et al., 2001). Two more decades of enterprise software investment have not made that vulnerability disappear. A later bibliometric review of the performance measurement literature found the field shifting toward more sophisticated, technology-enabled approaches, yet data integration and standardization challenges kept showing up as persistent, unresolved themes across the studies reviewed (Mishra et al., 2018).

In practice, fulfillment data for a typical enterprise sits scattered across enterprise resource planning systems, warehouse management platforms, transportation management systems, delivery and proof-of-delivery records, and logistics partner data that may never touch the enterprise's own systems at all. When these sources stay disconnected, nobody can easily trace an order from acceptance through allocation, picking, shipping, delivery, and proof-of-delivery capture as one continuous sequence. That is not just inconvenient. It leaves the organization structurally unable to tell a planning failure apart from a warehouse failure apart from a carrier failure, which happens to be exactly the distinction that matters when someone has to fix it.

2.2 *Control towers and governed KPI architectures*

The supply chain control tower concept is one response to this fragmentation, and recent work on control tower reference architectures shows how the underlying design tension actually plays out. Harmelink, Topan, and van Hillegersberg's (2025) reference architecture for service control towers in high-value asset maintenance names a dilemma that recurs across control tower designs: centralizing visibility for coordinated decision-making pulls against preserving the operational autonomy of the individual functions being monitored. Their architecture treats a control tower not as a single dashboard but as a layered decision-support structure, with data, coordination, and decision layers each governed on their own terms (Harmelink et al., 2025). That layered framing maps directly onto fulfillment intelligence. A milestone-based OTIF architecture runs into the same tension: milestone definitions need enough central governance to produce comparable metrics, while the systems generating those milestone events stay owned by different regional and functional teams who did not sign up to be centrally managed.

2.3 *Root-cause analytics in supply chain and manufacturing systems*

Related but separate literature addresses root-cause and bottleneck detection more directly. Lai, Shui, Ding, and Ni's (2021) work on data-driven dynamic bottleneck detection in complex manufacturing systems shows that the bottleneck's location can shift over time within a single production system, and that a static, one-time bottleneck analysis falls short for systems whose constraint points move with demand and configuration changes. Their setting is manufacturing production rather than order fulfillment, but the underlying argument transfers cleanly. A fulfillment lifecycle with more than thirty possible milestone checkpoints will not have one fixed point of failure, and any architecture built on a static assumption about where delays originate will lose relevance the moment operating conditions shift. Fulfillment intelligence architectures need the same dynamic, event-level granularity that dynamic bottleneck detection requires something in manufacturing, not a fixed diagnostic checklist someone wrote once and never revisited.

Regulated industries pose another level of difficulty. In pharmaceutical supply chains, for instance, Tucker and Daskin (2022) note the relationship between delivery reliability and drug shortage risk. Reliability failures in pharmaceutical supply chains carry consequences well beyond customer dissatisfaction and reach into patient access and continuity of care. Recent research from a worldwide benchmarking study of pharmaceutical manufacturing quality practices supports this reasoning. The maturity of an organization's quality management practices correlates with its delivery-related key performance indicators. This indicates that governance maturity and delivery performance move together instead of in opposition (Fellows et al., 2022). Taken together, the literature indicates that fulfillment intelligence in regulated supply chains is a governance and reliability question, rather than an operational efficiency question.

Regulated industries add another dimension. Tucker and Daskin's (2022) analysis of pharmaceutical supply chain reliability ties delivery reliability directly to drug shortage risk, arguing that reliability failures in pharmaceutical supply chains carry consequences well beyond customer dissatisfaction and reach into patient access and continuity of care. A global benchmarking study of pharmaceutical manufacturing quality practices found something consistent with this: the maturity of an organization's quality management practices correlates with its delivery-related key performance indicators, which suggests governance maturity and delivery performance move together rather than independently (Fellows et al., 2022). Taken together, this literature supports

treating fulfillment intelligence in regulated supply chains as a governance and reliability question first, and an operational efficiency question second.

3.0 Milestone-based fulfillment intelligence architecture

3.1 Lifecycle checkpoint design

The architecture proposed here treats each stage of the fulfillment lifecycle as a discrete, traceable data point, rather than folding the entire lifecycle into a single "delivered on time" flag. Representative checkpoints include order acceptance, inventory allocation, picking, dispatch, carrier handoff, delivery confirmation, and proof-of-delivery capture (Figure 1). A fully elaborated implementation can extend past thirty checkpoints once sub-stages within allocation, picking, and carrier transit get counted individually. Each checkpoint carries four attributes: an expected timestamp drawn from the standardized expectation model in Section 4, an actual timestamp captured from whichever source system owns that stage, a responsible party, and, when the checkpoint gets missed, an exception reason pulled from a governed taxonomy rather than free text someone typed in a hurry.



Figure 1. Fulfillment lifecycle milestone map: order acceptance, allocation, picking, dispatch, carrier handoff, delivery confirmation, and proof-of-delivery capture, each carrying an expected timestamp, actual timestamp, responsible party, and governed exception reason.

This checkpoint structure turns a single downstream OTIF metric into a sequence of comparisons between expected and actual timing at each stage. When a shipment misses its delivery window, the architecture does not just flag the miss. It identifies which checkpoint first drifted from its expected timing, and that is the piece of information an operations team can actually act on.

3.2 Data integration across planning, warehouse, and logistics systems

Populating this checkpoint structure means integrating data from enterprise planning systems, warehouse management platforms, transportation and carrier systems, and, where it exists, logistics partner data feeds. In a multi-region operating model spanning distinct planning and execution systems across geographies, this integration challenge gets compounded by inconsistent field definitions, time zone handling, and regional variation in how milestones get recorded in the first place. Cloud-based data platform architectures, built on governed ingestion and transformation layers using frameworks such as Databricks and comparable lakehouse environments, offer one practical way to reconcile these regional differences into a single, standardized event stream. The specific platform matters less than the governance discipline applied to it. Recent work on enterprise master data management makes a related point: data quality validation belongs at the point of integration, not after the fact, because that is what determines whether downstream analytics can be trusted at all (Korom, 2026). The same logic applies to milestone event integration. Expectation and actual timestamps that go unvalidated for consistency at ingestion will quietly corrupt every root-cause conclusion drawn from them later.

3.3 Governed common data objects

Because the source systems feeding this architecture were never designed to share a common data model, the architecture depends on a governed layer of common data objects defining what an "order," a "milestone," and an "exception" mean consistently across regions and business functions. Skip that governed layer, and dashboards built from different regional data sources will produce different answers to the same question, which undermines trust in the metric well before any root-cause analysis becomes possible.

This governance responsibility sits above any individual reporting tool. It has to be an architectural decision made once, not a configuration choice left to whichever team happens to build the next dashboard.

4.0 Standardized expectation modeling

4.1 Allocation, shipping, and delivery benchmarks

Milestone data only becomes useful once the organization has defined what "on time" means at every checkpoint, not only at final delivery. Allocation timing, shipping timing, delivery windows, and documentation timelines each need an explicit expectation, developed with the business functions that own the corresponding process, rather than inherited implicitly from whatever the source system happens to report. Gunasekaran and colleagues made a related point at the metric-design level: metrics chosen without reference to the process they measure tend to drift out of alignment with operational reality over time (Gunasekaran et al., 2001). The same risk shows up here at the checkpoint level. An allocation benchmark set without input from planning teams, or a delivery-window benchmark that ignores actual carrier transit variability by region, produces a metric that looks precise while measuring the wrong thing entirely.

4.2 Cross-regional governance of performance rules

In a global operating model, expectation modeling cannot be defined once centrally and applied uniformly, because carrier networks, regulatory documentation requirements, and typical transit times vary meaningfully across North America, Europe, Asia-Pacific, and Latin American operating regions. But expectation rules also cannot be left entirely to regional discretion, because that just reintroduces the inconsistency the architecture was built to eliminate. The practical resolution is a governed rule set with a common structure and regionally parameterized values. The same milestone definitions and exception taxonomy apply everywhere, while the specific timing thresholds attached to each milestone get calibrated by region and reviewed on a recurring governance cadence. This is an architectural responsibility, not a reporting configuration detail, and treating it as the latter is a common reason dashboards across regions end up disagreeing with each other.

5.0 Root-cause analysis and proactive bottleneck detection

5.1 From milestone signals to failure-mode diagnosis

Once milestone data is governed and integrated, root-cause analysis becomes a matter of comparing expected and actual timing across the checkpoint sequence for a given order or cohort of orders, instead of starting from a delivery failure and working backward through disconnected systems. Table 1 lays out representative on-time-in-full failure modes alongside the specific data needed to diagnose each one: allocation shortfalls traceable to inventory data, carrier delays traceable to transportation event data, documentation gaps traceable to proof-of-delivery capture systems. This structure moves the question from "Was this order late?" to "At which checkpoint did this order first fall behind its expected timing," which is a genuinely different and more actionable question to ask.

Table 1. Common On-Time In Full Delivery Failure Modes and the Data Needed to Diagnose Each One

Failure Mode	Likely Origin Stage	Diagnostic Data Required	Governing System of Record
Allocation shortfall	Inventory allocation	Available-to-promise balances, allocation timestamps, override logs	Enterprise planning system
Picking delay	Warehouse picking	Pick-task timestamps, labor and equipment availability records	Warehouse management platform
Carrier transit delay	Carrier handoff / transit	Carrier tracking events, transit-time benchmarks by lane	Transportation management system
Documentation gap	Proof-of-delivery capture	POD timestamp, partner submission logs, exception codes	Logistics partner / POD platform
Quantity shortfall	Order acceptance / allocation	Order line detail, fill-rate history, substitution records	Enterprise resource planning system

The dynamic bottleneck detection literature matters here too. Because the checkpoint most responsible for delays shifts with seasonal demand, carrier capacity, and regional operating conditions, a fulfillment intelligence architecture that only monitors a fixed set of "known" failure points will miss emerging bottlenecks until they become severe (Lai et al., 2021). A milestone architecture built to capture the full checkpoint sequence, rather than a curated subset chosen up front, stands a better chance of surfacing bottlenecks nobody anticipated when the system was first designed (Figure 2).

Figure 2. Centralized governed dashboard architecture

Order, logistics, proof-of-delivery, and partner data feed a governed common data object layer, which in turn supports milestone timing comparison, root-cause diagnosis, and executive KPI views.

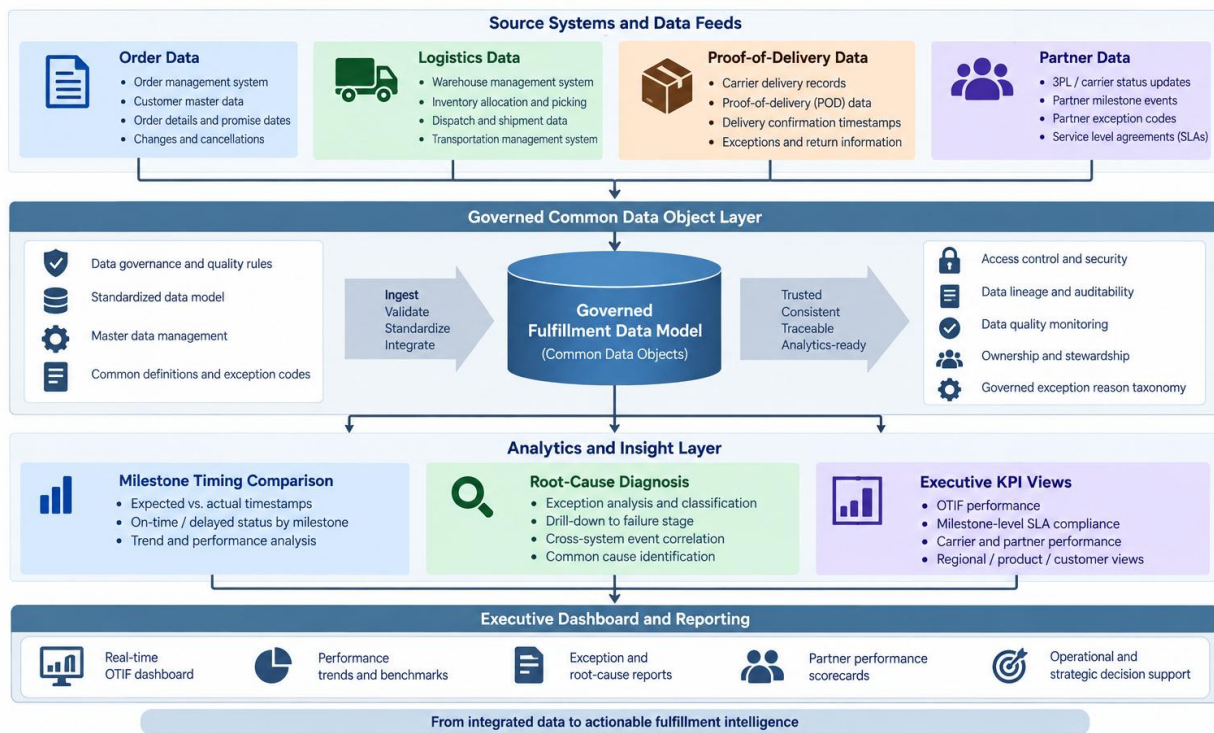


Figure 2. Centralized governed dashboard architecture: order, logistics, proof-of-delivery, and partner data feed a governed common data object layer, which in turn supports milestone timing comparison, root-cause diagnosis, and executive KPI views.

5.2 Partner collaboration and proof-of-delivery digitization

A substantial fraction of fulfillment failures occur outside of the enterprise system, in the partner systems used by carriers and logistics partners. Receiving proof-of-delivery, transit status, and exception notifications can be impacted by generation time delays and formats not suitable for an enterprise's systems. This is a common problem in the literature on freight and logistics technology. Integration pathways for digital platforms in road-based freight transport management between shippers, carriers, and logistics platforms are detailed by Heinbach et al. (2022). However, the maturity of integration for freight transport management platforms varies widely, even for powerful international logistic service providers. For example, in a milestone-based architecture, specification clarifies when the final milestones of the transportation lifecycle, delivery confirmation, and proof-of-delivery capture are presented as near-real time or ex post via manual reconciliation. Where partner integration stays immature, the architecture should make that gap explicit in its exception taxonomy rather than silently defaulting unavailable partner data to a result of on-time delivery, since that default would understate exactly the failure modes most likely to erode customer trust.

6.0 Partner collaboration and proof-of-delivery digitization

A meaningful share of fulfillment failures originate outside the enterprise's own systems entirely, in carrier and logistics partner operations that generate proof-of-delivery records, transit updates, and exception notifications on their own timelines and in their own formats. Digitizing and integrating this partner data is a well-documented challenge in the freight and logistics technology literature. Heinbach and colleagues' (2022) typology of digital platforms for road freight transport management describes the range of integration models connecting shippers, carriers, and logistics platforms and notes that integration maturity varies considerably even among large logistics providers. For a milestone-based architecture, partner data integration decides whether the final checkpoints in the lifecycle, delivery confirmation, and proof-of-delivery capture show up in near-real time or only get reconstructed after the fact through manual reconciliation. Where partner integration stays immature, the architecture should make that gap explicit in its exception taxonomy rather than quietly defaulting unavailable partner data to an on-time result, since that default would understate exactly the failure modes most likely to erode customer trust.

7.0 Discussion

The broader significance of this architecture reaches beyond operational efficiency. In global pharmaceutical and medical device supply chains, on-time and in-full delivery performance connects to patient access, continuity of care, and regulatory documentation obligations, not just to a customer satisfaction score (Fellows et al., 2022; Tucker & Daskin, 2022). A milestone-

based architecture that can show not only whether a delivery met its target but also which governed checkpoints it passed through and when also supports the audit and compliance documentation that regulated supply chains are expected to maintain.

Table 2 positions the milestone-based architecture against the two approaches most organizations already have in place: outcome-only OTIF reporting and a generic control-tower dashboard that centralizes visibility without necessarily governing the underlying milestone definitions. Broader editorial commentary on supply chain digitization has observed that digitization efforts across the field tend to advance data collection and visibility well ahead of the governance structures needed to make that data trustworthy across an organization (Tiwari et al., 2024). The comparison in Table 2 is meant to make that same gap concrete at the level of a single fulfillment architecture decision, rather than leave it as a general observation about the field.

Table 2. Comparative positioning of fulfillment visibility approaches

Dimension	Outcome-Only OTIF Reporting	Generic Control-Tower Dashboard	Milestone-Based Architecture (this article)
Failure traceability	Final result only	Partial, depends on source coverage	Checkpoint-level, full lifecycle
Root-cause speed	Manual, cross-system investigation	Faster, but limited by dashboard scope	Direct comparison of expected vs. actual per checkpoint
Governance model	Ad hoc, per-report	Centralized visibility, variable governance depth	Governed common data objects and regional rule calibration
Partner data handling	Often excluded	Inconsistent inclusion	Explicit exception taxonomy for partner gaps

8.0 Limitations

This architecture has real limits worth stating plainly. It does not report a validated numeric improvement in on-time-in-full rates, because that would require primary evaluation data this qualitative, architecture-focused contribution does not include. Any organization adopting this approach should plan its own before-and-after measurement rather than assume the gains described here transfer automatically. The governance model in Section 4 also assumes an organization has the cross-functional authority to define and enforce common milestone definitions across regions. That is an organizational precondition, not a purely technical one, and it is not guaranteed to exist everywhere. Finally, the architecture's ability to surface partner-side failure modes is bounded by the partner data integration maturity described in Section 6, so organizations with limited partner digitization will see reduced visibility into the lifecycle's final stages no matter how well the internal checkpoints are governed.

9.0 Conclusion

Treating On Time In Full delivery as a traceable lifecycle rather than a single terminal outcome changes what an organization can actually do when performance slips. A milestone-based architecture built on governed common data objects, standardized and regionally calibrated expectation models, and integrated planning, warehouse, and logistics data turns delivery reporting from a backward-looking scorecard into a structure that can point to the specific place where something went wrong. In regulated, multi-region supply chains, where delivery reliability carries consequences well beyond customer satisfaction, this shift from outcome reporting to governed fulfillment intelligence is not a reporting enhancement. It is closer to a foundational requirement for resilient operations.

Acknowledgments

The author used Claude (Anthropic) to assist with drafting and language refinement. The conceptual framework, its architectural design decisions, and all analytical conclusions originated with and were determined by the author, who takes full responsibility for the accuracy, originality, and integrity of this manuscript.

Funding

This research received no external funding.

Author Contribution

Venugopal Kakarla is the sole author and is responsible for conceptualization, methodology, investigation, writing of the original draft, writing review and editing, and visualization.

Conflict of Interest

The author declares no conflict of interest.

Data Availability

This article describes a system architecture and its design rationale, grounded in the cited literature. No proprietary or client-identifying data are disclosed, no primary numeric results are reported, and no new datasets were generated or analyzed.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

- Fellows, M., Friedli, T., Li, Y., Maguire, J., Rakala, N., Ritz, M., Bernasconi, M., Seiss, M., Stiber, N., Swatek, M., & Viehmann, A. (2022). Benchmarking the quality practices of global pharmaceutical manufacturing to advance supply chain resilience. *The AAPS Journal*, 24(6), Article 111. <https://doi.org/10.1208/s12248-022-00761-7>
- Gunasekaran, A., Patel, C., & Tirtiroglu, E. (2001). Performance measures and metrics in a supply chain environment. *International Journal of Operations & Production Management*, 21(1/2), 71-87. <https://doi.org/10.1108/01443570110358468>
- Harmelink, R., Topan, E., & van Hillegersberg, J. (2025). The service control tower: dilemmas, decisions and a reference architecture for maintenance of (high-value) assets. *Information Systems and e-Business Management*, 23(4), 931-969. <https://doi.org/10.1007/s10257-025-00705-6>
- Heinbach, C., Beinke, J., Kammler, F., & Thomas, O. (2022). Data-driven forwarding: a typology of digital platforms for road freight transport management. *Electronic Markets*, 32(2), 807-828. <https://doi.org/10.1007/s12525-022-00540-4>
- Korom, P. (2026). Data quality validation in enterprise master data management. *Proceedings of the 15th International Conference on Data Science, Technology and Applications*. <https://doi.org/10.5220/0015161700004091>
- Lai, X., Shui, H., Ding, D., & Ni, J. (2021). Data-driven dynamic bottleneck detection in complex manufacturing systems. *Journal of Manufacturing Systems*, 60, 662-675. <https://doi.org/10.1016/j.jmsy.2021.07.016>
- Mishra, D., Gunasekaran, A., Papadopoulos, T., & Dubey, R. (2018). Supply chain performance measures and metrics: a bibliometric study. *Benchmarking: An International Journal*, 25(3), 932-967. <https://doi.org/10.1108/bij-08-2017-0224>
- Tiwari, M. K., Bidanda, B., Geunes, J., Fernandes, K., & Dolgui, A. (2024). Supply chain digitisation and management. *International Journal of Production Research*, 62(8), 2918-2926. <https://doi.org/10.1080/00207543.2024.2316476>
- Tucker, E. L., & Daskin, M. S. (2022). Pharmaceutical supply chain reliability and effects on drug shortages. *Computers & Industrial Engineering*, 169, Article 108258. <https://doi.org/10.1016/j.cie.2022.108258>

Biographical Sketch

Venugopal Kakarla is a data architect with over ten years of experience in enterprise data engineering, specializing in Azure Databricks-based supply chain and healthcare and medical technology data platforms, including on-time-in-full delivery visibility, order-risk monitoring, and available-to-promise prediction across global operating regions. He holds Microsoft Azure certifications (DP-203 Azure Data Engineer, DP-204 Azure Developer, and DP-305 Azure Solution Architect) and a B.Tech. in Computer Science and Engineering from Vignan's Lara Institute of Technology & Science.