Unified Namespace: The Scalable Foundation for *AI in Manufacturing*

Capgemini | snowflake®

Executive Summary

The manufacturing sector stands at a critical inflection point. The promise of Artificial Intelligence (AI) to catalyze the fourth industrial revolution, unlocking unprecedented efficiency, resilience, and innovation is immense. Yet, for a vast majority of organizations, this promise remains unfulfilled. Most AI initiatives are stuck in "pilot purgatory," failing to scale beyond isolated experiments to deliver enterprise-wide business value. The root cause of this stagnation is not a deficiency in AI technology or a lack of data, but a foundational crisis in data architecture. The modern factory, while generating terabytes of data daily, is crippled by fragmented data silos, a deep cultural and technical chasm between Information Technology (IT) and Operational Technology (OT), and an unmanageable web of brittle, point-to-point integrations. This fractured landscape creates an environment where AI simply cannot thrive or scale.

This whitepaper presents a definitive, three-part solution to this AI paradox, outlining a synergistic approach that combines a modern architecture, a powerful platform, and an expert partner to build a truly intelligent enterprise.

01

A Modern Architecture: The Unified Namespace (UNS). The foundational layer is the Unified Namespace, an architectural concept that establishes a single source of truth or all industrial data. Built upon an event-driven message broker, the UNS decouples disparate systems, from the shop floor to enterprise IT. It provides a standardized, contextualized stream of real-time data, organized according to industry standards like ISA-95, making information universally accessible and understandable. This architecture is the antidote to data silos and integration complexity.

02

The Mind of an UNS Architecture: The Snowflake® AI Data Cloud. The enterprise core of this solution is the Snowflake AI Data Cloud. Snowflake is purpose-built to ingest the unified, real-time data streams from the UNS and converge them with all other enterprise data sources (e.g., ERP, SCM, CRM). This breaks down silos at a macro level, creating a single, governed repository for the entire business. More than just a data warehouse, Snowflake provides a fully integrated environment to build, deploy, operationalize, and govern AI and Machine Learning (ML) models at scale, directly on top of this live, trusted data foundation.

03

A Realization and Integration Partner: Capgemini. Technology and architecture alone are insufficient. The journey to an intelligent enterprise is a profound business transformation that requires deep domain knowledge and expert guidance. Capgemini, with its Intelligent Industry vision and proven transformation methodologies, serves as the integrator. Capgemini provides the critical end-to-end services to architect the solution, integrate the technology stack, and most importantly manage the significant organizational and cultural changes necessary to bridge the IT/OT divide and ensure lasting success.

This synergistic partnership between a revolutionary architecture (UNS), a dominant AI data platform (Snowflake), and a world-class transformation partner (Capgemini) offers a de-risked, accelerated, and value-driven path for manufacturers. It is the blueprint for moving beyond experimentation to build a resilient, adaptive, and truly intelligent enterprise, powered by data and AI.

Table of contents

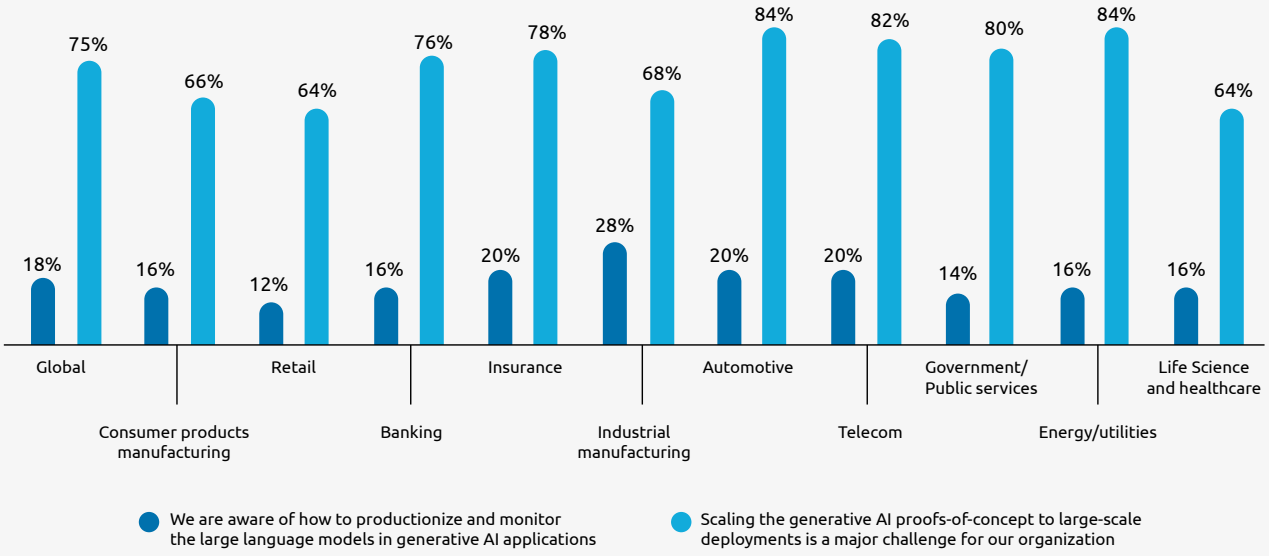
01

The AI Paradox in Manufacturing: Drowning in Data, Starving for Insight

The ambition to create smarter, more efficient, and more resilient manufacturing operations through digital transformation is universal. Industry 4.0 and Artificial Intelligence are no longer futuristic concepts but immediate strategic imperatives. However, a persistent and costly gap exists between this ambition and the reality of execution.

1.1 The unfulfilled promise of Industry 4.0 and AI

The promise of AI in manufacturing is immense: predictive maintenance, optimized supply chains, automated quality control, and hyper-personalized products. Recognizing this, organizations are investing heavily in AI and generative AI capabilities. However, progress has been slower than anticipated. While 60% of organizations have launched AI pilots, a staggering 75% of data executives cite the large-scale deployment of these proofs of concept as a major challenge.



Source: Capgemini Research Institute, Data-powered enterprises survey, April 2024, N = 500 organizations represented by 500 data executives and 500 business executives.

Our research reveals that nearly three in five executives recognize that, in order to employ generative AI to its full potential, there is a need for significant alterations to data collection, storage, retrieval, reusability, and governance. Much data remains trapped in silos, with only half of organizations possessing clear processes for integrating data across functions. Only 51% of data executives say that their organization has clear processes to manage siloed data and data integration across functions. Moreover, 49% of data sources are in the cloud, with the remainder still residing in local servers, posing accessibility challenges. Moreover, most organizations (87%) have yet to utilize external data sources.

This creates the central paradox of modern manufacturing: facilities are generating unprecedented volumes of data, with a single modern factory capable of producing one terabyte per day, yet they are struggling to make this data fit for purpose to extract its true value. The result is a landscape littered with promising but stalled proofs-of-concept, where organizations are drowning in a sea of raw data but starving for the actionable insights that AI promises to deliver.

1.2 The root cause: A fractured data landscape

The failure to scale AI in manufacturing is not an AI problem; it is a data architecture and organizational problem. The foundation upon which these advanced analytics are meant to be built is fundamentally broken, fractured by deep-seated technical and cultural divisions.

Data Silos

The most immediate problem is data fragmentation. For decades, manufacturing data has been segregated into distinct, non-communicating silos. Production data, real-time sensor readings, machine states, process parameters are locked away in Operational Technology (OT) systems on the shop floor, such as SCADA and historians. Meanwhile, business data, production schedules, inventory levels, customer orders, and financial information, resides in Information Technology (IT) systems like Enterprise Resource Planning (ERP) and Manufacturing Execution Systems (MES). This separation, often exacerbated by legacy ERP systems designed as independent, transaction-oriented applications, makes it impossible to achieve a holistic, end-to-end view of operations. Without the ability to correlate shop floor events with business context, any meaningful, enterprise-level analysis is severely handicapped.

Why point-to-point integrations fail at scale

In an attempt to bridge the IT/OT divide, the default approach has been to create direct, point-to-point API connections between specific applications. An engineer might write a custom script to pull data from a historian and push it to an analytics dashboard. While this may solve an immediate, isolated problem, it is a strategy that is brittle, unscalable, and ultimately exacerbates the core issues.

This method creates a “spaghetti architecture” - a complex, invisible, and unmanageable web of hard-coded integrations. These connections are difficult to maintain, document, and secure. As factory equipment is upgraded or production programs are changed, a frequent occurrence in modern manufacturing, these brittle integrations inevitably break, leading to undetected bad or missing data for weeks or months. This approach also leads to what can be called “model purgatory,” where data architects get stuck in an endless and futile effort to create a single, perfect, all encompassing logical data model of the entire factory. In reality, no single end application can consume such a complex model, and the effort only serves to delay value and increase complexity.

This architectural approach is the very antithesis of the agility required for modern AI development. It reinforces data silos and creates a self-perpetuating cycle of organizational friction and poor technical choices. Any solution that aims to scale AI successfully must first dismantle this flawed foundation. It requires a new architectural paradigm that decouples systems to solve the technical problem, and a new operational model that bridges the organizational divide to solve the human problem.

02

The Architectural Blueprint: Introducing the Unified Namespace (UNS)

To break the cycle of data fragmentation and integration fragility, manufacturers need to adopt a new architectural paradigm. The Unified Namespace (UNS) provides this modern blueprint. It is not a single product or technology, but a strategic approach to data architecture designed to create a single, reliable, and universally accessible source of truth for all industrial data.

2.1 Defining the UNS: A single source of truth for the modern enterprise

A Unified Namespace is a consolidated, abstracted data structure through which all business applications from the plant floor to the executive suite can consume real-time industrial data in a consistent and standardized manner. The core objective of the UNS is to eliminate data silos by centralizing data access into a common namespace, organized using a common hierarchy and data model.

Instead of applications polling thousands of individual devices for data, the UNS operates on an event-driven publish-subscribe model. Data-producing systems (e.g., PLCs, sensors) publish their data to a central location, and data-consuming systems (e.g., MES, ERP, AI platforms) subscribe to the specific data streams they need. This architecture is designed to be lightweight, open, and highly scalable, fundamentally shifting the data-sharing paradigm from slow, cyclical requests to efficient, real-time push communication.

2.2 The building blocks: Data modeling, contextualization, and ISA-95

The creation of a functional and scalable UNS rests on three foundational pillars: rigorous data modeling, comprehensive contextualization, and a standardized hierarchical structure.

Data modeling is essential

At its core, a UNS is built from well-defined data models. A data model is a standardized, rich description of a physical or logical thing, such as a pump, a production line, or a work order that consolidates raw data points with their essential context to address a specific business use case. Effective data modeling is critical because it enables:

Standardization: It defines a consistent way to categorize and structure data across the enterprise.

Interoperability: It allows users and systems from different functions (e.g., operations, maintenance, quality) to look at the same data and immediately understand its source, structure, and meaning.

Data Governance: It provides a framework for dictating how information should be shared and who is authorized to access it.

A common pitfall is to create models that are overly complex from the start. The best practice is to begin with a specific, critical use case and model only the minimum data required, iterating and expanding the model over time as new requirements emerge.

Models vs. Instances

A key concept for achieving standardization at scale is the distinction between a model and an instance. The model serves as the generalized, reusable blueprint for a type of asset or process. For example, an organization might create a single "Industrial Tightening System" model that defines the standard attributes for all its tightening tools, such as TargetTorque, FinalAngle, Status, and LastResult.

An instance, on the other hand, is a unique, digital representation of a specific real-world asset. Using the Industrial Tightening System model, engineers would create ten distinct instances to represent ten physical screwdrivers on the assembly line (e.g., TighteningSystem_A5_S12, Screwdriver_RobotArm_3, etc.). Each instance is then populated

with the connections and data tags that pull live, real-time values from its corresponding physical tool. This approach ensures that while every tightening system reports its data in a standardized format defined by the model, each instance contains the unique values for that specific asset. This prevents instance-specific information (like "TighteningSystem_A5_S12 FinalAngle") from being hard-coded into the reusable model, a common mistake that severely limits scalability.

The role of ISA-95

To organize these models and instances in a logical and navigable way, the UNS leverages the ISA-95 international standard for enterprise-control system integration. ISA-95 defines a standard hierarchical model for manufacturing operations, breaking the enterprise down into logical levels: Enterprise > Site > Area > Production Line > Work Cell.

This hierarchy provides the perfect structure for the UNS topic namespace. For example, data for a specific machine might be published under the topic MyCompany/SpringfieldPlant/Assembly/Line1/TighteningSystem_A5_S12. This structure makes the data self-organizing and allows consuming applications to easily browse and subscribe to data at any level of the hierarchy, from an entire site down to a single sensor.

Contextualization

The final and most crucial step is contextualization. This is the process of transforming raw, cryptic OT data into valuable, self-describing information. This is achieved by including metadata directly within the data model itself. Instead of just publishing a raw value, the model includes attributes for its UnitOfMeasure, AssetID, PhysicalLocation, MinMaxRange, and other descriptive data. When this fully contextualized payload is published to the UNS, any consuming application receives not just the what (the data value) but also the so what (its meaning and context), eliminating the ambiguity and tribal knowledge that plagues traditional data integration efforts.

2.3 The role of the event-driven message broker

The architectural linchpin that makes the Unified Namespace possible is the central message broker. A UNS cannot function without a broker; it is the software component that acts as the traffic cop for all data, receiving every published message and routing it to all subscribed applications.

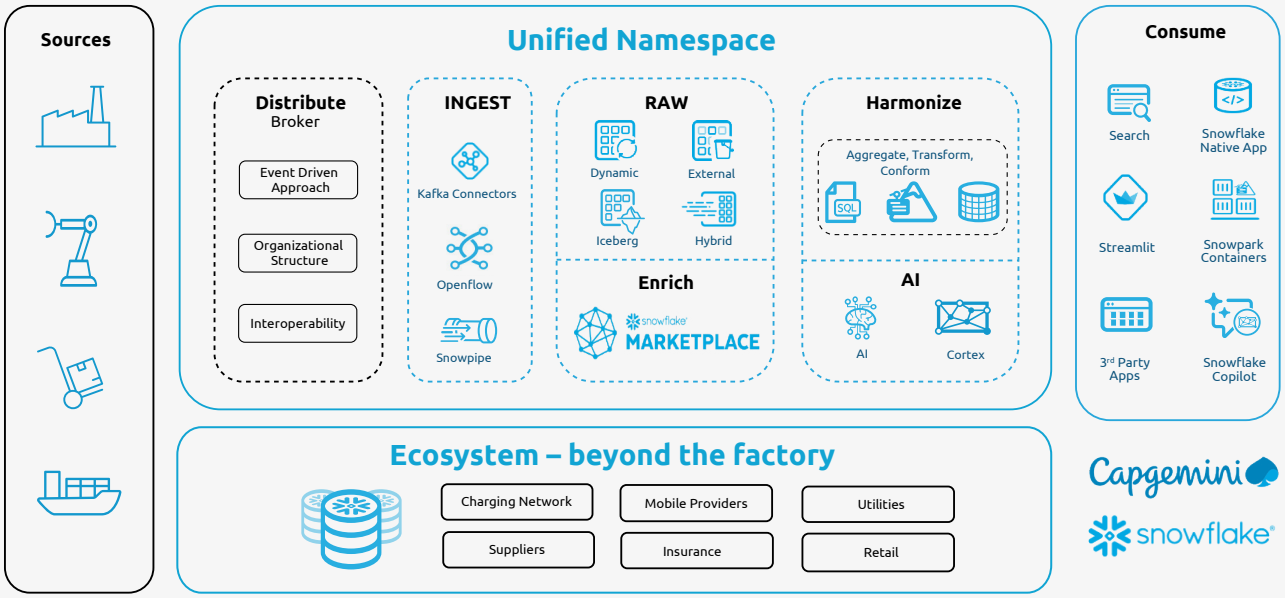
The single most important function of the broker is to decouple data producers from data consumers. In a UNS architecture, a PLC on the factory floor publishes its data to the broker without needing to know or care if one application or one hundred applications will consume it. Likewise, an AI platform in the cloud can subscribe to that data stream without needing to establish a direct, fragile connection to the PLC itself. This loose coupling is the antidote to the brittle "spaghetti architecture" of point-to-point integrations.

This technical decoupling has profound business implications. It fosters immense organizational agility. When the data science team requires a new data point for an AI model, they simply subscribe to a new data topic on the broker. They do not need to file a request with the OT team to engineer a new direct connection, a process that could take months and require scheduled downtime. This dramatically shortens innovation cycles and allows IT, OT, and data science teams to evolve their respective systems independently and in parallel, directly mitigating the cultural friction and project delays identified as a core industry problem.

2.4 The missing puzzle piece - The Mind

“When I first began architecting and evangelizing the Unified Namespace concept back in 2019, there was always one missing puzzle piece that has become increasingly essential: the Mind. To simplify the concept, I often compare the UNS architecture to the human body. We have our senses: sight, hearing, touch, which are the equivalent of industrial sensors, constantly gathering data from the surrounding environment. Our nervous system acts as the central data distributor, the broker, transmitting this sensory data, which is an extremely important step. But the data is meaningless without our mind. We still require our memory and cognitive functions to refine the information, apply more context, and define actions, which are then sent back through the nervous system. This ‘Mind’ is the missing piece, and it is critically important for any UNS architecture that aims to generate AI potential at scale.”

- Marian Raphael Demme, Principal - Capgemini



03

Powering the UNS: The Snowflake AI Data Cloud as the Enterprise Data Core

While the Unified Namespace provides the revolutionary architecture for collecting, contextualizing, and sharing real-time operational data, a second, equally powerful layer is required to turn that data into enterprise-wide business value. The Snowflake AI Data Cloud serves as this layer the destination for unified data and the engine for global-scale analytics and AI.

3.1 From shop floor to top floor: Ingesting and governing unified data

The broker excels at unifying data at the edge and plant level. Snowflake excels at consolidating this unified data with all other enterprise data sources from ERP, SCM, and CRM systems at a global scale. Snowflake's AI Data Cloud for Manufacturing offers a single, fully managed, multi-cloud platform designed to break down these macro-level data silos and create a single source of truth for the entire business.

A key technical advantage is Snowflake's native support for semi-structured data formats like JSON, which is the standard encoding for UNS payloads published via OPC UA over MQTT. This allows Snowflake to ingest the rich, contextualized data streams from the UNS broker seamlessly, without requiring complex and costly Extract, Transform, Load (ETL) processes that often strip away context. The data lands in Snowflake with its structure and metadata intact, ready for immediate use. This provides a single, governed platform where all enterprise data is consolidated, secured by robust access policies, and made available in near real-time to any authorized user or application across clouds.

3.2 Unleashing AI: Developing and deploying models directly on governed data

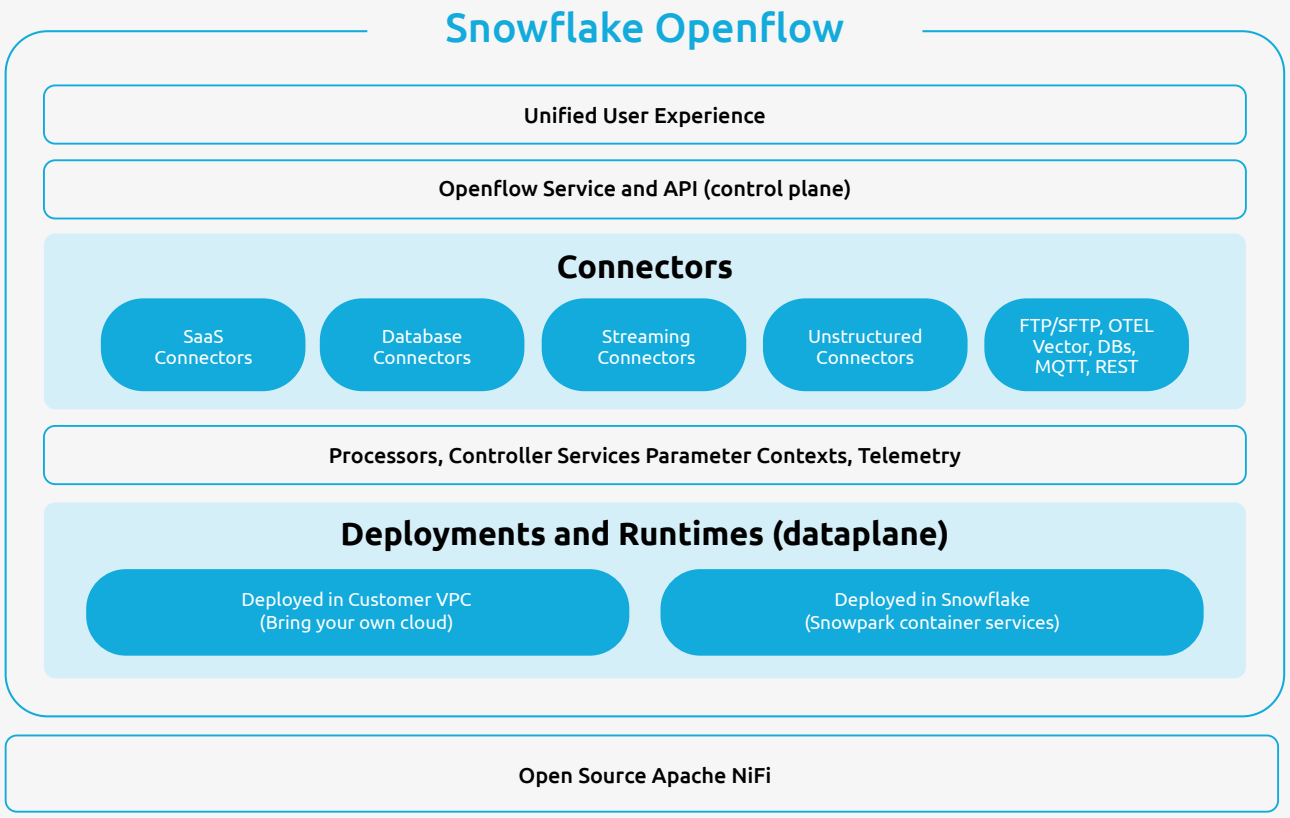
Snowflake is far more than a data warehouse; it is an end-to-end platform for the entire AI and machine learning (ML) lifecycle. This integrated nature solves one of the biggest challenges in enterprise AI: the common and risky practice of moving data out of the governed corporate platform to separate, specialized tools for model development. By bringing the compute to the data, Snowflake ensures that AI models are built, trained, and deployed on the most current, complete, and governed enterprise data set, which is a critical factor for model accuracy, reliability, and trust.

Snowflake ML and Snowpark: These integrated sets of capabilities empower data scientists and ML engineers to accelerate their workflows dramatically. Using *Snowflake Notebooks*, they can perform large-scale development and distributed training directly within Snowflake's secure perimeter.

With Snowpark, they can write complex data transformations and model training logic in familiar languages like Python, which is then executed on Snowflake's powerful and elastic compute engine, including GPU-powered instances for demanding deep learning workloads. This streamlined process has been shown to deliver significant results; for example, the loyalty program Scene+ used Snowflake ML to cut its model time-to-production by over 60% and reduce costs by over 35%.

Snowflake Cortex AI: For organizations looking to leverage the power of generative AI, Cortex AI provides a suite of serverless, ready-to-use AI functions and access to industry-leading large language models (LLMs). This service simplifies and democratizes the creation of AI applications. A prime example is Siemens Energy, which used Cortex AI to build a document Q&A chatbot.

Openflow: A managed integration service that accelerates secure ingestion and transformation at scale. It provides curated connectors, streaming and batch pipelines, and advanced text processing (e.g., semantic chunking) to operationalize AI workloads. It delivers governance, lineage, enterprise observability, and data quality controls.



A major hurdle in AI is operationalization, or MLOps the process of reliably deploying and maintaining models in production. Many models that perform well in a data scientist's lab fail in the real world due to issues like data drift, complex deployment pipelines, or a lack of monitoring. Snowflake solves this last mile problem by providing an integrated solution. Models can be deployed as scalable endpoints using

Snowpark Container Services, and their performance and data quality can be continuously monitored using built-in ML Observability tools. When this robust MLOps framework is fed by the consistent, standardized, real-time data stream from a UNS, it creates an incredibly resilient end-to-end pipeline from sensor to production AI, drastically reducing the risk and complexity of enterprise AI initiatives.

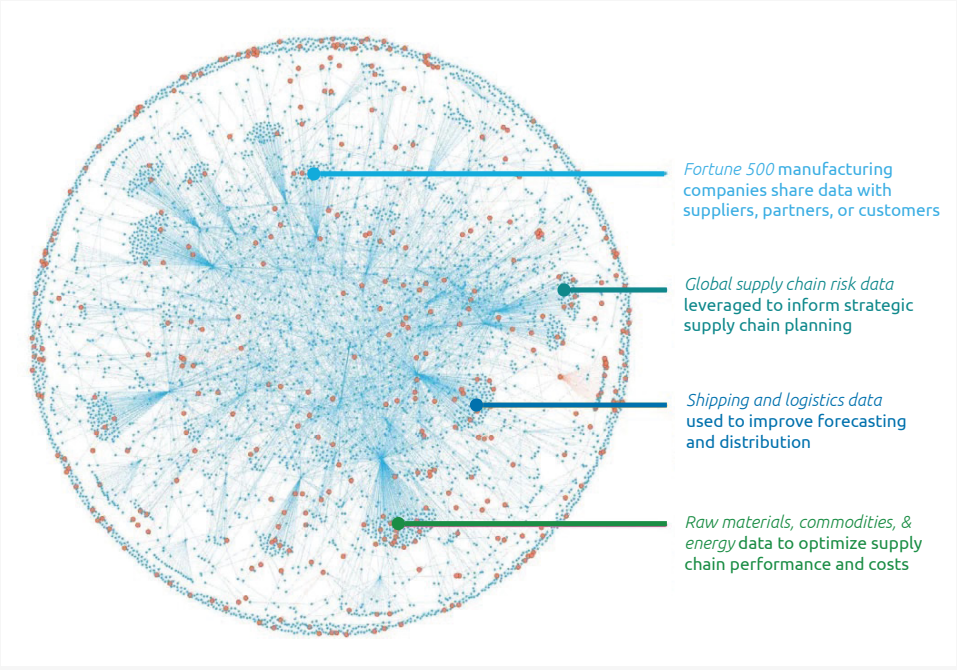
3.3 Beyond the factory: Secure collaboration across the entire value chain

Modern manufacturing is not a solitary endeavor; it is a collaborative ecosystem of suppliers, partners, logistics providers, and customers. Success depends on the seamless and secure flow of information across this value chain.

Snowflake's unique secure data sharing capabilities are a game-changer in this regard. Without physically copying or moving data a process that is slow, expensive, and creates security risks, manufacturers can provide live, read-only, and highly governed access to specific slices of their data with ecosystem partners. For instance, a manufacturer can grant a key supplier real-time visibility into inventory levels of a critical component, allowing the supplier to proactively manage their own production and shipping schedules. Similarly, they can share detailed quality and traceability data with customers, enhancing trust and transparency. This capability transforms the supply chain from a series of disconnected, reactive links into a connected, proactive, and resilient network.

Snowflake's AI Data Cloud for Manufacturing

Thousands of organizations across industries collaborating with their ecosystem



* Visualization based on actual Data Cloud sharing activity as of July 31, 2023

04

Powering AI and Analytics: The Snowflake AI Data Cloud

Snowflake's AI Data Cloud offers manufacturers a robust and versatile platform that directly addresses their unique data challenges and unlocks the potential of Industry 4.0 and AI. Here are key areas of interest:

Unified data ingestion and consolidation from OT and IT:

Manufacturers grapple with fragmented data, with Operational Technology (OT) data (sensor readings, machine states, process parameters) siloed from Information Technology (IT) data (ERP, SCM, CRM, MES). Snowflake directly solves this by providing a single, governed repository that ingests and converges real-time data streams from a Unified Namespace (UNS) with all other enterprise data sources. This capability is crucial for achieving a holistic, end-to-end view of operations, which is often impossible with traditional fragmented data landscapes.

1. Native support for industrial data formats: A significant advantage for manufacturers is Snowflake's native support for semi-structured data formats like JSON. This is particularly relevant as JSON is the standard encoding for UNS payloads published via OPC UA over MQTT, allowing Snowflake to seamlessly ingest rich, contextualized data streams from the UNS broker without complex and costly ETL processes that might strip away vital context. The data lands in Snowflake with its structure and metadata intact, ready for immediate use.

2. Scalability for terabytes of industrial data: Modern factories generate enormous volumes of data, with a single factory capable of producing one terabyte per day. Snowflake's multi-cloud platform is designed to handle this scale, ensuring that manufacturers can store, process, and analyze vast amounts of operational and business data without performance bottlenecks, supporting deep historical analysis and real-time insights.

Advanced analytics and machine learning for operational excellence:

Snowflake is not just a data warehouse; it's an end-to-end platform for the entire AI and machine learning lifecycle. This is critical for manufacturers aiming for predictive maintenance, optimized supply chains, automated quality control, and hyper-personalized products.

1. In-Platform ML development with Snowpark: Data scientists and ML engineers can leverage Snowflake ML and Snowpark to build, train, and deploy AI models directly within Snowflake's secure perimeter using familiar languages like Python. This eliminates the common and risky practice of moving data out of the governed corporate platform to separate tools for model development, ensuring models are built on the most current, complete, and governed enterprise data. This streamlined process has been shown to significantly cut model time-to-production and reduce costs.

2. MLOps for reliable production AI: Many promising AI models fail in the real world due to operationalization challenges. Snowflake addresses this last mile problem by providing an integrated solution for deploying models as scalable endpoints using Snowpark Container Services and continuously monitoring their performance and data quality with built-in ML Observability tools. This robust MLOps framework, fed by the consistent, real-time data from a UNS, creates a resilient end-to-end pipeline from sensor to production AI, drastically reducing the risk and complexity of enterprise AI initiatives.

3. Generative AI with Cortex AI for innovation: For leveraging generative AI, Snowflake Cortex AI provides serverless, ready-to-use AI functions and access to industry-leading large language models (LLMs). This democratizes the creation of AI driven applications, enabling manufacturers to build solutions like document Q&A chatbots from vast archives of paper records, transforming them into searchable, actionable data and deriving insights faster.

Secure data sharing across the value chain:

Modern manufacturing thrives on collaboration across suppliers, partners, logistics providers, and customers. Snowflake's secure data sharing capabilities are a game-changer in this regard. Manufacturers can provide live, read-only, and highly governed access to specific data slices with ecosystem partners without physically copying or moving data, which is slow, expensive, and risky. This allows for:

1. Near real-time supplier visibility: Granting a key supplier real-time visibility into inventory levels of critical components, enabling proactive management of their own production and shipping schedules.

2. Enhanced customer transparency: Sharing detailed quality and traceability data with customers, fostering trust and transparency.

3. Building a resilient supply network: Transforming the supply chain from disconnected, reactive links into a connected, proactive, and resilient network.

05

From Blueprint to Reality: Capgemini as Transformation Partner

A robust architecture and a powerful platform are necessary but insufficient for success. The transition to an AI driven enterprise is a complex business transformation that requires a partner with a rare combination of deep technical expertise, extensive manufacturing domain knowledge, and a proven methodology for managing organizational change. Capgemini is this indispensable partner, providing the strategic guidance and hands-on execution needed to turn the blueprint into a reality.

Building a scalable and secure UNS is a highly specialized task. Capgemini possesses deep, demonstrated technical expertise in designing and implementing these architectures, leveraging open standards to ensure maximum interoperability and future-proofing.

Capgemini's consultants understand the critical nuances of selecting the right message broker, defining optimal payload encoding formats, and structuring data models to meet specific business needs.

This expertise is centralized within Capgemini's Intelligent Industry area dedicated to helping clients navigate the convergence of the digital and engineering worldsthe very challenge at the heart of this transformation. Their approach is not merely to connect systems, but to partner with clients to devise, design, and deliver their holistic vision for a smart, connected, and sustainable operation. This is supported by end-to-end data strategy and factory services, which establish the solid foundation of data governance, management, and quality assurance required for both the UNS and its integration with the Snowflake AI Data Cloud.

06

The Joint Value Proposition: Building Your AI-Ready Future, Today

The combined offering from Capgemini and Snowflake, built upon the architectural foundation of the Unified Namespace, provides a clear and synergistic pathway for manufacturers to accelerate their journey toward becoming an intelligent enterprise. This is best illustrated by examining a common, high-value use case and summarizing the transformative shift in a clear framework.

The manufacturing industry is at a crossroads. The path of incremental, siloed technology adoption has led to the current AI paradox, where vast data assets fail to translate into tangible business value. The alternative path, the one that leads to a truly intelligent enterprise, requires a more deliberate and holistic strategy. It is a journey that necessitates a fundamental transformation of architecture, platforms, and processes.

This whitepaper has laid out a clear blueprint for this transformation. It begins with a modern architectural foundation, the Unified Namespace, that decouples systems and creates a single, contextualized source of truth for all operational data. This foundation feeds a powerful enterprise core, the Snowflake AI Data Cloud, where OT and IT data are converged, and where AI and ML models can be built, deployed, and governed at scale on live, trusted data. Guiding this entire journey is an expert partner Capgemini that provides the rare blend of technical, domain, and change management expertise required to navigate the complexities of this transformation and ensure its success.

About Capgemini

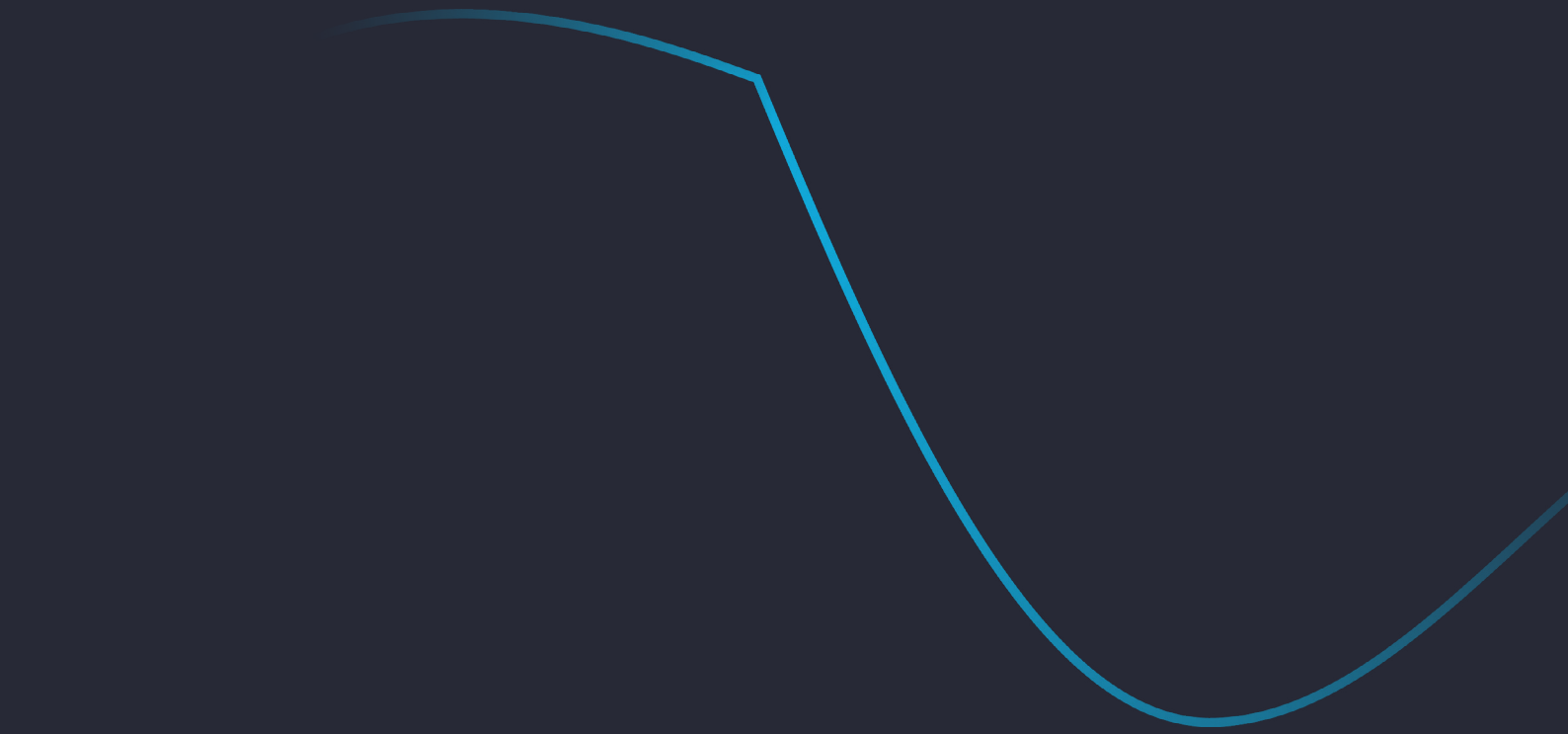
Capgemini is a global business and technology transformation partner, helping organizations to accelerate their dual transition to a digital and sustainable world, while creating tangible impact for enterprises and society. It is a responsible and diverse group of 350,000 team members in more than 50 countries. With its strong over 55-year heritage, Capgemini is trusted by its clients to unlock the value of technology to address the entire breadth of their business needs. It delivers end-to-end services and solutions leveraging strengths from strategy and design to engineering, all fuelled by its market leading capabilities in AI, generative AI, cloud and data, combined with its deep industry expertise and partner ecosystem. The Group reported 2024 global revenues of €22.1 billion.

Get the future you want | www.capgemini.com

About Snowflake

Snowflake is the platform for the AI era, making it easy for enterprises to innovate faster and get more value from data. More than 12,000 customers around the globe, including hundreds of the world's largest companies, use Snowflake's AI Data Cloud to build, use and share data, applications and AI. With Snowflake, data and AI are transformative for everyone.

Learn more at www.snowflake.com (NYSE: SNOW)



Capgemini