INDIUM

Agentic Data Platforms: From AI-Ready Data to Autonomous Enterprise

A Practical Enterprise Framework for Data Products, Knowledge Products, Semantic Intelligence, and AI Agents



Why the next wave of enterprise AI will be won on data foundations, not model sophistication – and how leaders can move from AI experimentation to autonomous operations at scale.

Research paper



Ayushi Jain
Manager- Analyst Relations
Author



Ashish Kumar
Chief AI Architect
Reviewer

Contents

1. Executive Summary....3

2. Introduction: The Enterprise AI Paradox....4

3. Three Decades of Enterprise Data Architecture....7

4. Why Traditional Data Architectures Are No Longer Enough....7

4. AI-Ready Data Architecture: A Leap Toward Efficiency....8

4.1. What changes between traditional and AI-ready data....8

4.2 Business outcomes AI-ready data delivers....9

5. The Anatomy of an AI-Ready Data Architecture....10

6. The Emergence of Knowledge Products....11

6.1 Ontologies: the vocabulary that gives data its meaning....11

7. The AI-Ready Data Architecture Maturity Model....12

8. The Agentic Data Platform....13

8.1 Core capabilities....14

8.2 Who needs it most....14

8.3 Assessing readiness before you build....15

9. The Enterprise Transformation Roadmap....15

10. Market Evidence: From Data Architecture to Agentic Data Platforms....16

11. Adoption Challenges and How to Overcome Them....18

12. Key Questions for Enterprise Leaders....19

13. What the Future Looks Like for Visionary Leaders....13

14. Conclusion....21

15. References....23

© Indium

Executive Summary

Agentic AI – systems that perceive context, reason, plan, and autonomously execute tasks toward business goals – is the next evolution of enterprise artificial intelligence. Yet the gap between ambition and outcome is stark. MIT research found that 95% of enterprise generative AI pilots deliver no measurable profit-and-loss impact despite an estimated US$30–40 billion of investment[3], and Gartner predicts that over 40% of agentic AI projects will be canceled by the end of 2027 due to escalating costs, unclear business value, or inadequate risk controls[1].

This paper argues that the primary barrier is not the maturity of AI models. It is the underlying data architecture. Enterprise data estates were designed for analytics and human consumption; they lack the business context, semantics, governance, and lineage that AI agents need to interpret information and make trustworthy autonomous decisions. Accenture finds that only 7% of enterprises have data foundations required to scale advanced AI[4].

The paper introduces an **Agentic Data Platform:** an AI-native data architecture that connects governed data, business context, semantic intelligence, knowledge products, and AI agents so that systems can discover information, reason over it, and execute business processes with appropriate governance and human oversight.

The paper presents a five-layer anatomy of AI-ready data products, a six-level maturity model, a seven-phase transformation roadmap, and showcases relevance through concrete market evidence.

Who should read this: Chief data, analytics, information, and technology officers; enterprise and data architects; and business leaders sponsoring AI transformation programs. No specialist data engineering background is assumed.

© Indium

1. Introduction: The Enterprise AI Paradox

Unlike traditional automation, which executes predefined rules, agentic AI continuously perceives, reasons, learns, collaborates, and acts toward defined business goals within governance and oversight boundaries. Advances in large language models (LLMs), retrieval-augmented generation (RAG), and reasoning engines have made increasingly capable enterprise applications possible – from self-service data integration and anomaly detection to automated decision support.

But today, adoption is racing ahead of readiness. According to a McKinsey report, 88% of organizations now use AI in at least one business function, yet only about a third have begun scaling it enterprise-wide[13] and only 17% of organizations have deployed AI agents to date, while more than 60% plan to within two years[9]. S&P Global finds that the average organization scrapped 46% of AI proofs-of-concept before they reached production[5].

This is called the Enterprise AI Paradox: organizations continue to invest heavily in increasingly sophisticated AI capabilities while the architectures underneath remain designed for reporting, dashboards, and human interpretation. AI systems can reason only as well as the data and knowledge available to them, and today, the semantic layer is widely recognized as the essential infrastructure for analytics. However, fragmented data, inconsistent semantics, missing business context, and governance gaps – not model weakness – prevent AI from understanding enterprise intent and producing reliable, actionable outcomes.

Critical Business Challenges

As enterprises move from AI experimentation to scaled adoption, several structural challenges continue to limit their ability to realize sustained business value.

01 Scalability

AI initiatives may demonstrate strong results in controlled pilots but struggle to scale across business units, geographies, and complex enterprise environments, exposing gaps in data, infrastructure, governance, and operating models.

02 Fragmented Data

Data distributed across warehouses, lakehouses, applications, and departmental repositories creates silos, making it difficult for AI systems to access a complete and consistent view of enterprise information.

03 Legacy Modernization

Enterprises often struggle to connect modern AI applications with legacy systems and architectures, limiting the flow of data and preventing AI from being embedded into critical business workflows.

04 Limited Architectural Visibility

Technology leaders often lack a comprehensive understanding of the capabilities, dependencies, limitations, and AI-readiness of their existing technology and data environments. As a result, organizations may invest in new AI capabilities without addressing the underlying constraints that prevent them from scaling.

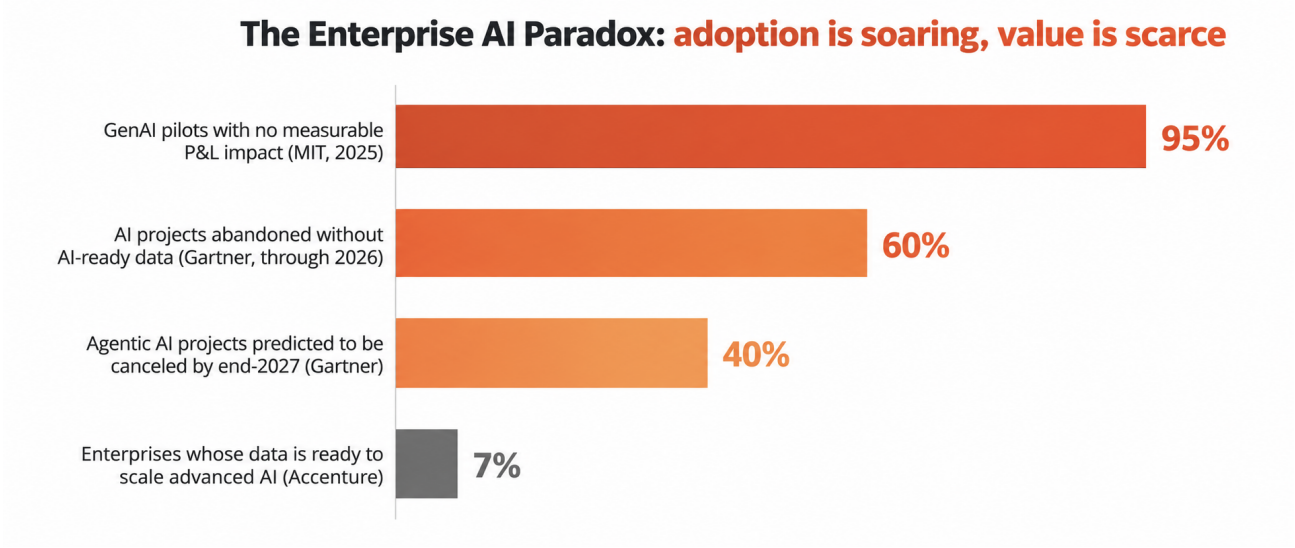


Figure 1. Enterprise AI investment is surging, but structural data gaps keep value out of reach. Sources: MIT NANDA (2025); Gartner (2025); Accenture (2026).

Closing this gap requires a new enterprise data paradigm – one that moves beyond making data accessible toward making enterprise knowledge contextual, governed, machine-readable, and actionable. The remainder of this paper maps that journey.

© Indium

2. Three Decades of Enterprise Data Architecture

Each generation of data architecture solved a specific business problem while exposing a new limitation that drove the next wave of innovation. The evolution has now reached an inflection point: the primary consumer of enterprise information is no longer a human analyst but an intelligent AI agent.

Architecture	What it solved	Why it was not enough
Operational databases	Reliable storage of day-to-day transactional data.	Unsuitable for enterprise-wide analytics and decision support.
Data warehouses	Centralized, structured data for reporting and historical analysis.	Limited support for unstructured data, real-time analytics, and exploding volumes.
Data lakes	Any data, any format, at scale, for exploration and machine learning.	Weak governance and metadata created untrusted "data swamps."
Lakehouses	Lake scale with warehouse governance and performance.	Technically unified but infrastructure-centric; no business context or domain ownership.
Data mesh	Decentralized, domain-owned data shared as a product.	Establishes ownership but not the machine-readable semantics AI systems need.
Data products	Reusable, governed data assets enabling self-service analytics.	Designed for dashboards, reports, and human interpretation — not autonomous agents.
Knowledge products	Data fused with semantics, rules, lineage, and decision logic that AI can reason over.	Individually powerful, but still need a platform to orchestrate discovery, reasoning, and execution.
Agentic Data Platform	Unified ecosystem orchestrating knowledge products, semantic services, governance, and AI agents.	The foundational architecture for the autonomous enterprise.
Autonomous enterprise	AI agents continuously perceive, reason, and execute; humans provide strategic oversight and exception management.	The desired end state of enterprise AI transformation.

Although enterprise data architecture has evolved significantly, the transformation has rarely been linear. New paradigms have emerged alongside existing ones, leaving enterprises with a complex mix of databases, data warehouses, data lakes, and lakehouses.

The same pattern emerged in the 1990s and 2000s, when BI platforms such as Cognos, MicroStrategy, and Hyperion offered tightly coupled analytics ecosystems that delivered value but also created data silos due to limited enterprise integration.

© Indium

Even as enterprises explore more advanced approaches to AI-ready and knowledge-driven architectures, many continue to operate data lakes and lakehouses. This is not necessarily a reflection of technological inertia; in many cases, organizations face practical barriers that make migration

Three factors largely explain this persistence:

01

Limited Migration Imperative

Existing platforms may continue to meet current analytical and operational requirements, reducing the immediate business justification for migration.

02

Migration Cost and Complexity

Moving large volumes of enterprise data, applications, pipelines,and dependencies to a new architecture can require significant financial investment, technical effort, and operational disruption.

03

Sunk Investments and Ecosystem Dependencies

Organizations that have already invested heavily in platforms, tools, skills, and integrations are often reluctant to replace them, particularly when critical business processes depend on the existing ecosystem.

3. Why Traditional Data Architectures Are No Longer Enough

Data products marked a genuine leap: they treated data as a reusable, governed, domain-oriented business asset and democratized access to trusted information. But their consumers were business analysts, data scientists, and reporting applications. They expose what the data contains – not what it means, why it matters, how it should be interpreted, or how it should influence a decision. Autonomous agents need exactly those missing ingredients.

The cost of this gap is not theoretical. Gartner estimates poor data quality alone costs the average organization US$12.9 million per year6, before counting the AI initiatives it silently kills. Six limitations recur across enterprises:

- **Limited business context.** Datasets, metrics, and APIs are exposed without business intent, organizational policy, or decision rationale, so agents misinterpret what they retrieve.
- **Absent semantic intelligence.** Schemas and technical metadata are not business semantics. Without ontologies, knowledge graphs, and shared vocabularies, agents cannot infer relationships or reason across domains.

© Indium

- **Human-centric design.** Dashboards and SQL assume a human will interpret the result. Agents require machine-readable knowledge that can be consumed without human mediation.
- **Knowledge trapped outside data.** Policies, business rules, regulatory requirements, and standard operating procedures live in documents, emails, and employee expertise, leaving AI with an incomplete picture.
- **No support for autonomous reasoning.** Causal relationships, dependencies, constraints, confidence measures, and historical decision patterns are rarely represented explicitly.
- **Governance built for data, not for AI.** Agentic AI adds new requirements – explainability, reasoning transparency, policy-aware decisions, confidence scoring, and continuous validation of AI-generated actions.

The implication is direct: enterprises must move beyond engineering reusable data assets toward engineering reusable knowledge assets. That shift – from data-centric to knowledge-centric architecture – is the foundation of everything that follows.

4. AI-Ready Data Architecture: A Leap Toward Efficiency

AI-ready data is trusted, governed, context-rich enterprise data that AI systems can consume directly to deliver intelligent, secure, compliant, and scalable business outcomes. Unlike traditional datasets, it is engineered for machine understanding: high-quality data combined with business context, semantic meaning, metadata, governance, and domain knowledge.

4.1 What changes between traditional and AI-ready data

Dimension	Traditional data (built for humans)	AI-ready data (built for humans and machines)
Question answered	"What happened?" — for analysts and BI tools.	"What happened, why, so what, and what next?" — for intelligent agents.
Structure	Complete, accurate records in tables and fields.	Records plus explicit relationships, hierarchies, patterns, and rules within and across datasets.
Context	Technical metadata and standard values.	Semantics, ontologies, worked examples, and domain-specific business meaning.
Freshness	Periodic batch refresh for reporting.	Real-time, so systems see operational change as it occurs.
Governance	Ownership and auditability of data.	Ownership plus policy-aware, permission-aware, explainable AI consumption.
Purpose	Reporting and analytics.	Action and better decision-making.

4.2 Business outcomes AI-ready data deliver

- **Accelerated AI adoption.** Ready-to-use, trusted datasets shorten the path from pilot to production – directly influencing and accelerating the adoption rate.

- **Reduced operational cost.** Less time and effort are spent on data preparation for analytics, reporting, and model development, and less of the US$12.9 million average annual leakage from poor data quality[6], thereby increasing operational efficiency.

- **Better business decisions.** Models trained and grounded on reliable, governed, compliant data produce answers leaders can defend to boards and regulators.

- **Secure data collaboration.** Governed, permission-aware sharing across business units, industries, and geographies fosters innovation within AI ecosystems without losing control.

To become truly AI-ready, organizations must integrate knowledge from structured and unstructured sources – enterprise documents, process models, diagrams, source code, policies, knowledge bases, and institutional expertise – capturing not only business facts but the relationships, terminology, and decision logic that let AI reason effectively.

5. The Anatomy of an AI-Ready Data Product

An AI-ready data product is not a single dataset. It is a layered architecture in which each layer adds intelligence, trust, and usability – and each answers a distinct question.

© Indium

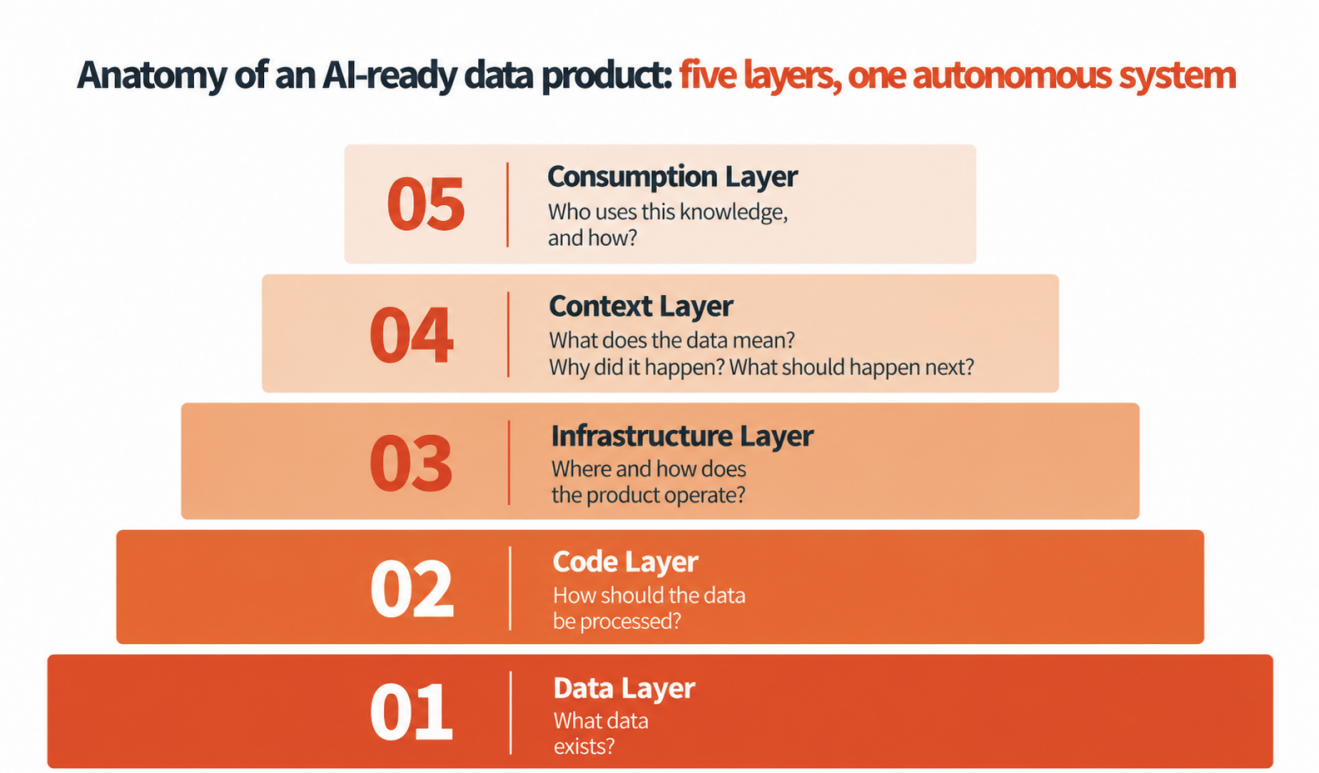


Figure 2. The five layers of an AI-ready data product. The context layer is what separate it from a traditional data product.

Layer	Role	Contains
1. Data	The core. Raw enterprise data generated across business systems.	Transactional records, operational tables, documents, logs, images, IoT events, real-time streams.
2. Code	The logic. Turns raw data into usable, quality-assured information.	ETL/ELT pipelines, transformation logic, validation rules, data quality tests, governance policies.
3. Infrastructure	The execution. Makes the product scalable, secure, and continuously available.	Cloud storage, compute engines, orchestration, catalogs, pipelines.
4. Context	The intelligence. The defining layer — enriches information with business meaning machines can read.	Reasoning engines integrated with business metadata, semantic definitions, ontologies, lineage, knowledge graphs, enterprise policies.
5. Consumption	The experience. Enables humans and agents to consume knowledge through interfaces and services.	Dashboards, APIs, AI copilots and agents, enterprise applications, multi-agent systems.

With the reasoning layer in place, AI moves beyond answering what happened to explaining why it happened, what it means, and what should happen next – and why that is the recommended course of action.

© Indium

6. The Emergence of Knowledge Products

A knowledge product is not a document repository, an intranet library, a semantic layer, or a collection of prompts. It is a structured, reusable, governed, machine-readable representation of enterprise knowledge that combines trusted data with business context, semantics, rules, policies, relationships, and decision logic – enabling humans and AI systems to interact efficiently. It is the bridge between data and decisions: decision intelligence, packaged for reuse.

Worked example: credit risk in banking

A bank's 200-page markdown lending policy document contains knowledge — but it is not a knowledge product. A **credit-risk knowledge product** packages borrower information, transaction history, credit reports, risk indicators, lending policies, regulatory requirements, business definitions, entity relationships, historical decisions, and reasoning context into one governed, machine-readable asset.

In practice, that asset is typically a versioned YAML or JSON manifest (or its knowledge-graph equivalent) that an agent retrieves directly.

With it, an AI system can predict and analyze risk, explain its assessment, recommend actions consistent with policy, and trigger the appropriate workflow — with every step traceable for regulators.

6.1 Ontologies: the vocabulary that gives data its meaning

Enterprise data arrives from many sources in many forms, and the same business term often carries different meanings across departments. A sales team may build meticulous tables and dashboards, yet a new analyst – or an AI agent – still cannot tell where customer data lives or how entities relate, because meaning was never made explicit. Without a shared understanding of business concepts, AI produces inaccurate responses, inconsistent insights, and unreliable decisions.

An ontology solves this by establishing a standardized semantic framework: the structured vocabulary that defines business entities (Customer, Product, Order), standardized metric definitions (Revenue, Customer Lifetime Value, Churn), relationships ("Customer places Order"), and the business rules and constraints that govern interpretation. Realized at scale as a knowledge graph, it lets AI agents reason consistently across fragmented systems.

© Indium

The evidence: context is measurable accuracy

In a peer-reviewed enterprise benchmark, a GPT-4-class model answering business questions directly over a SQL database achieved 16.7% accuracy — and **0% on schema-intensive questions** about metrics and KPIs. The same model over a knowledge graph representation of the same database achieved 54.2% — a **3x improvement**[7]. A follow-up study adding ontology-based query checking and automated repair raised the gain to **4.2x**[8].

Enterprises typically build this as three connected layers on top of raw structured data: **semantic models** (the business glossary and metric definitions), the **ontology** (entities, relationships, and rules), and the **knowledge graph** (the ontology instantiated at scale across live data) — together forming the context layer that enabled the models to reason accurately.

The lesson for leaders: investing in semantics is not an academic nicety. It is the must-to-have accuracy investment for enterprise AI.

Building an ontology is an ongoing business discipline, not a one-time technical exercise. Enterprises typically combine two approaches: top-down, in which leaders and domain experts define entities, relationships, and constraints (every new dataset or metric gets a definition, related entities, governing rules, and interpretive context – not just a name); and bottom-up, in which vocabulary is extracted from existing databases, documents, and feeds such as customer reviews that are integrated into the data flow. The result is a trusted foundation on which humans and agents can interpret enterprise data identically, regardless of source or format.

Reference architecture: The three context layers above form a consistent stack on top of raw structured data, feeding the knowledge products and agents described in Sections 6 and 8: Structured data → Semantic models (glossary, metric definitions) → Ontology (entities, relationships, rules) → Knowledge graph (the ontology instantiated at scale) → Knowledge products (Section 6) → AI agents operating through the Agentic Data Platform (Section 8). Each layer is what makes the one above it machine-consumable – this is the architecture that produced the accuracy gains cited above.

7. The AI-Ready Data Architecture Maturity Model

Most enterprises cannot leap directly to autonomy; each level below builds the capabilities the next level assumes. Leaders should locate their organization honestly – most sit at Levels 2 or 3 today – and invest in the specific gap to the next level rather than in generic "AI transformation." The Agentic Data Platform, introduced in Section 8, is what operationalizes Levels 4–6 at enterprise scale – it is the platform layer that lets knowledge products and AI agents function together, not a separate initiative from this model.

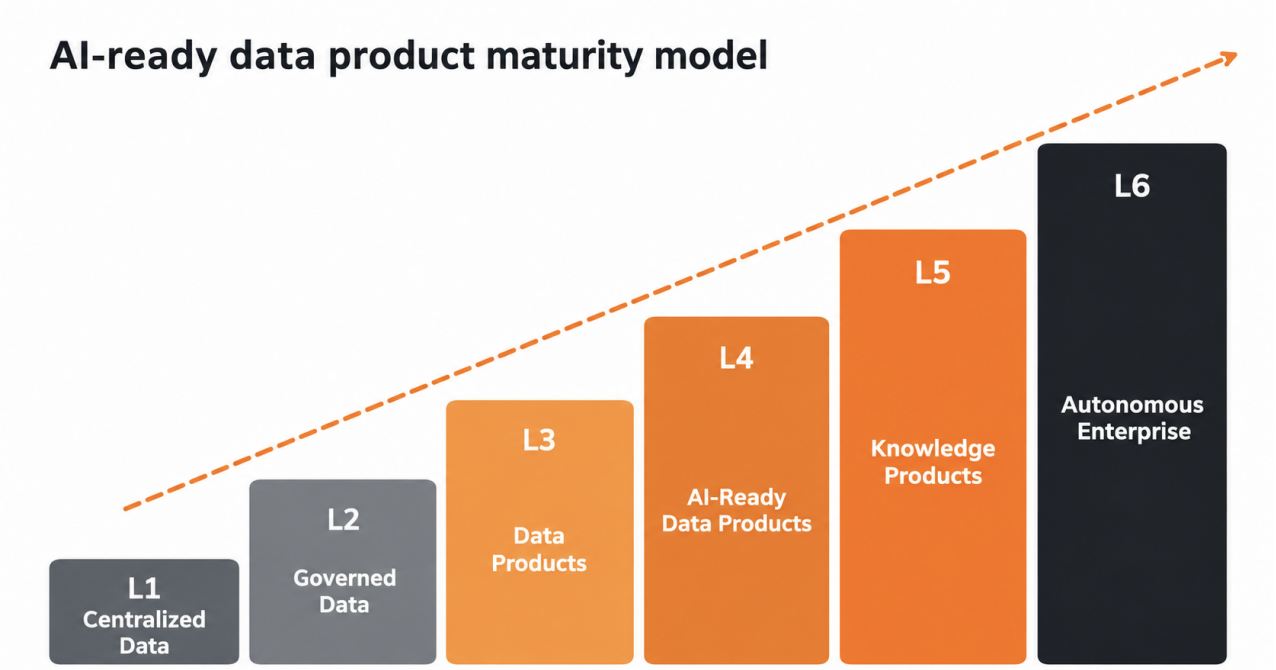


Figure 3. Six levels of enterprise capability, from centralized data to the autonomous enterprise.

Level	Enterprise capability	What it unlocks
Level 1	Centralized data	A single place to find enterprise data; basic reporting.
Level 2	Governed data	Trusted, quality-controlled, access-managed data; auditability and compliance.
Level 3	Data products	Domain-owned, reusable assets; self-service analytics at scale.
Level 4	AI-ready data	Context-rich, machine-readable data that AI systems can consume safely and accurately.
Level 5	Knowledge products	Packaged business knowledge — semantics, rules, and reasoning context — powering explainable AI decisions.
Level 6	Autonomous enterprise	AI agents execute end-to-end processes with humans providing oversight and exception management.

8. The Agentic Data Platform

Individual knowledge products provide domain intelligence, but enterprises still need an integrated platform to orchestrate discovery, reasoning, collaboration, and autonomous execution across domains. The Agentic Data Platform is an AI-native, context-aware ecosystem that unifies data across disparate systems, synchronizes it in real time, standardizes metadata, preserves business context, and enriches everything with semantic intelligence – so agents retrieve the right information, reason accurately, and execute workflows with confidence.

© Indium

8.1 Core capabilities

- **Unified data fabric.** A real-time fabric connecting data across business functions, applications, and clouds – eliminating silos and giving agents and humans one consistent, up-to-date view of enterprise information.

- **Contextual intelligence.** Data enriched with metadata, ontologies, business rules, and governance policies, making AI-generated insights accurate, explainable, and compliant by construction.

- **Multi-agent orchestration.** Specialized agents – ingestion, quality, governance, security, reasoning, analytics, workflow execution – collaborating and exchanging context to optimize complex processes, with resilience no single model can offer.

- **Self-learning and optimization.** Continuous learning from interactions, patterns, feedback, and outcomes: detecting anomalies, flagging emerging risks, recommending corrective actions, and refining reasoning over time.

8.2 Who needs it most

Agentic Data Platforms deliver the greatest value to organizations that are scaling generative and agentic AI from pilots to enterprise-wide deployment; modernizing complex hybrid and multi-cloud estates; eliminating data silos across functions; unifying a governed view of enterprise knowledge for cross-functional collaboration; and building AI-ready data products for trusted, explainable, autonomous decision-making.

© Indium

8.3 Assessing readiness before you build

Readiness extends beyond technology to people, data, governance, and integration maturity. Three dimensions provide a practical assessment framework:

Dimension	Core question	Key considerations
Organizational readiness	Do our people and operating model support AI-driven systems?	AI and data literacy across business and IT; executive sponsorship and change management; data ownership and stewardship; trusted, governed data assets.
Context foundation	Can AI understand the business behind our data?	Standardized metadata and business glossary; enterprise ontologies and semantic models; business rules and process logic; governance, compliance, and policy frameworks.
Integration readiness	Can systems exchange information securely and consistently?	API-first architecture; connectivity across cloud, on-premises, and SaaS; secure data sharing and privacy controls; semantic interoperability across systems.

Evaluating these dimensions before implementation surfaces capability gaps early, reduces deployment risk, and establishes the foundation to scale agentic initiatives successfully.

9. The Enterprise Transformation Roadmap

The roadmap below lets enterprises locate their current position, understand their limiting constraint, and plan the move to the next phase. The phases are sequential in dependency but can overlap in execution; the semantic work of Phases 3–4 is where most programs stall, and where the accuracy dividends documented in Section 6 are earned.

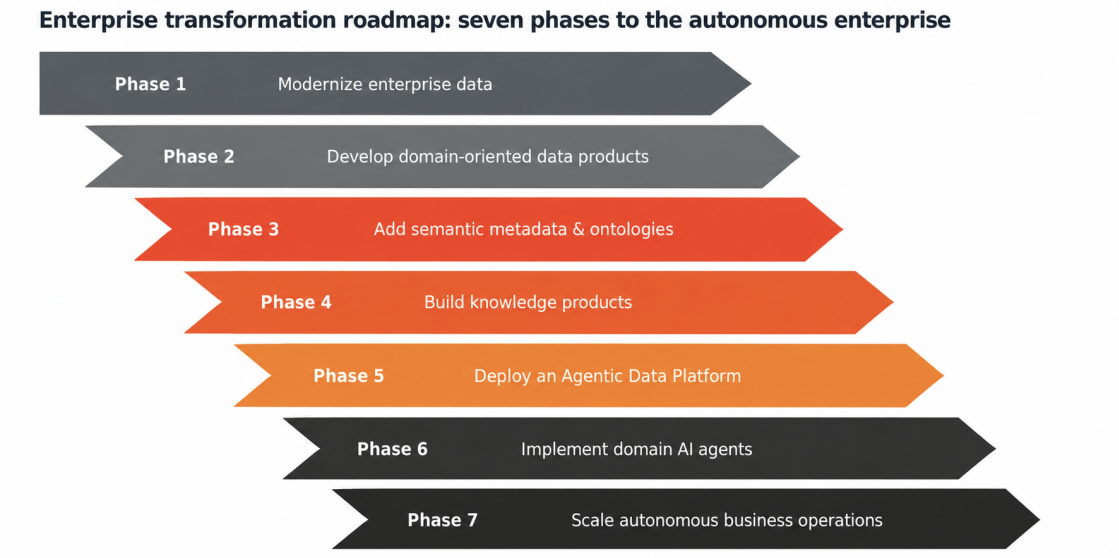


Figure 4. Seven phases from modernized data to autonomous operations.

Phase	Focus	Exit criteria (what "done" looks like)
01 Modernize enterprise data	Consolidate estates onto governed cloud/Lakehouse foundations; automate quality.	Critical domains catalogued; quality measured; access governed.
02 Develop domain data products	Assign domain ownership; publish reusable, discoverable data products.	Named owners; SLAs; self-service consumption in priority domains.
03 Add semantic metadata	Define business glossary, metrics, ontologies; connect scattered sources via knowledge graph.	Shared vocabulary in force; key entities and relationships machine-readable.
04 Build knowledge products	Package data + context + rules + reasoning context per business problem.	First knowledge products powering explainable AI decisions in production.
05 Deploy an Agentic Data Platform	Unify fabric, semantics, governance, and agent services into one platform.	Agents access governed, real-time, permission-aware enterprise knowledge.
06 Implement domain AI agents	Launch specialized agents for priority workflows with human oversight.	Measured outcomes per agent; escalation and audit paths proven.
07 Scale autonomous operations	Extend multi-agent orchestration across functions; humans manage exceptions.	End-to-end processes run autonomously within governance boundaries.

10. Market Evidence: From Data Architecture to Agentic Data Platforms

The transition toward Agentic AI is increasingly influencing the evolution of enterprise data platforms. While the market is still developing and the concept of an Agentic Data Platform is not yet standardized, leading technology providers are introducing capabilities that point toward a common architectural direction.

Recent announcements from major cloud and data platform providers illustrate this shift. Rather than treating data platforms as passive repositories for analytics and reporting, these platforms are increasingly being designed to provide AI agents with business context, semantic meaning, lineage, governed access, and the ability to execute actions.

PROVIDER
Snowflake

EMERGING CAPABILITY
Snowflake CoWork

CONTEXT & KNOWLEDGE LAYER

Semantic views, governed context, metadata, schemas, and enterprise data provide agents with business-relevant information. Snowflake also positions Horizon Context and Semantic Views as foundations for trusted, business-ready data.

AGENTIC CAPABILITY

CoWork supports analytical search, deep research, data exploration, automation, and agent-based interaction with enterprise data.

RELEVANCE TO THE AI-READY DATA PARADIGM

Demonstrates the movement from traditional data access toward context-aware AI consumption of enterprise data.

PROVIDER
Databricks

EMERGING CAPABILITY
Databricks Genie Ontology

CONTEXT & KNOWLEDGE LAYER

The context layer is created through a map of snippets that Genie extracts from existing assets, including metric views, dashboards, SQL queries, and Genie agents.

AGENTIC CAPABILITY

Genie Ontology adds a business-aware map of the enterprises that is automatically extracted and maintained from existing assets, providing modeled context that is defined, governed, and certified.

RELEVANCE TO THE AI-READY DATA PARADIGM

Reflects the proposed shift toward an AI-native data architecture built around unified systems wrapped with context, semantics, governance, and autonomous action.

PROVIDER
Google Cloud

EMERGING CAPABILITY
Agentic Data Cloud

CONTEXT & KNOWLEDGE LAYER

Knowledge Catalog aggregates metadata and enriches it with business semantics, relationships, lineage, and contextual information to create a governed context layer for agents.

AGENTIC CAPABILITY

Enables agents to perceive, reason, and act across enterprise data through an AI-native data architecture, including data agents and Deep Research capabilities.

RELEVANCE TO THE AI-READY DATA PARADIGM

Closely reflects the proposed shift toward an AI-native data architecture built around context, semantics, governance, and autonomous action.

© Indium

Implication for the Enterprises

The emerging capabilities from Databricks, Snowflake and Google Cloud do not represent a single standardized architecture, but they demonstrate a clear industry direction: enterprise AI requires more than models and data access. It requires a governed layer of business context and knowledge.

This reinforces the central thesis of this paper: the next evolution of enterprise data is not simply about creating better data products, but about transforming data products into reusable Knowledge Products that can serve as the foundation for Agentic AI and the Autonomous Enterprise.

11. Adoption Challenges and How to Overcome Them

Most enterprises know what they want; the real challenge is figuring out how to achieve it. The barriers below recur across industries – and each has a practical countermeasure rooted in the roadmap.

According to Everest Group, 23% of organizations cite legacy IT systems and siloed data architectures as challenges while 18% report insufficient tools to handle unstructured data volumes & complexities.

Barrier	Why it hurts	Countermeasure
Legacy architecture	Agents cannot reach systems of record; integration disrupts workflows and inflates cost.	Phase 1 modernization; API-first integration; wrap legacy systems with governed access services rather than rip-and-replace.
Limited metadata access	Without lineage and definitions, AI outputs cannot be trusted or explained.	Automate metadata capture in pipelines; make catalog entries a release requirement for every data product.
Fragmented systems	Knowledge scattered across silos forces agents to act on partial pictures.	Unified data fabric with real-time synchronization; knowledge graph to connect structured, unstructured, and raw sources.
Poor data quality & visibility	US$12.9M average annual cost and a leading cause of abandoned AI projects[2].	Automated quality gates and continuous monitoring; quality metrics tied to the maturity model, reported to leadership.
Governance complexity & regulation	Agentic AI adds explainability, transparency, and validation duties on top of privacy and compliance obligations.	Policy-as-code in the context layer; permission-aware access; audit trails and confidence scoring for every AI action.
Undefined value & "agent washing"	Hype-driven projects without ROI discipline drive the predicted 40% cancellation rate; Gartner estimates only ~130 of thousands of self-described agentic vendors are genuinely agentic[1].	Pursue agentic AI only where clear value or ROI is defined; baseline metrics before deployment; rethink workflows rather than bolting agents onto old ones.

12. Key Questions for Enterprise Leaders

1. Is AI-ready data the same as high-quality data?

No. High-quality data is accurate, complete, consistent, and reliable. AI-ready data goes further by adding the context, semantics, metadata, relationships, governance, and business meaning required for AI systems to interpret and reason over information effectively.

2. Do enterprises need to replace their existing data platforms to become AI-ready?

Not necessarily. AI readiness should be viewed as an evolutionary journey rather than a complete technology replacement. Enterprises can progressively enrich existing data warehouses, lakehouses, and other platforms with semantic, contextual, governance, and knowledge capabilities.

3. Is a Knowledge Product simply a semantic layer or knowledge graph?

No. Semantic layers and knowledge graphs can be components of a Knowledge Product, but they do not constitute one by themselves. A Knowledge Product brings together data, semantics, business context, policies, domain knowledge, decision logic, governance, and reasoning context into a reusable business asset.

4. Why can't enterprises simply give AI agents access to their existing data?

Because access does not equal understanding.

Enterprise data may be fragmented across systems, contain inconsistent definitions, lack business context, or exist without the relationships and policies required to interpret it correctly. AI agents need more than data retrieval—they need contextual knowledge to reason reliably and take appropriate action.

© Indium

5. Do Knowledge Products replace data products?

No. They build upon them. Data Products establish trusted, governed, reusable data assets. Knowledge Products add the contextual, semantic, and decision intelligence required to make that data actionable for AI systems.

6. How does Agentic AI benefit from Knowledge Products?

Knowledge Products provide AI agents with the business context and reasoning foundation required to move beyond information retrieval. They enable agents to interpret enterprise data, apply business rules and policies, understand relationships, evaluate alternatives, and recommend or execute appropriate actions within defined governance boundaries.

7. Can organizations build Knowledge Products without migrating their existing data?

Yes. Knowledge Products can be built by connecting and enriching existing enterprise data sources rather than requiring wholesale migration. This enables organizations to preserve existing technology investments while progressively introducing AI-ready capabilities.

8. How should enterprises begin their journey toward an Agentic Data Platform?

Enterprises should begin by assessing their data, architecture, governance, and AI readiness. Rather than attempting enterprise-wide transformation immediately, organizations can identify high-value business domains, establish trusted data products, enrich them with context and semantics, develop Knowledge Products, and progressively introduce agentic capabilities.

9. What is the ultimate objective of an Agentic Data Platform?

The objective is not to create another data platform. It is to create an enterprise foundation in which trusted data can be transformed into contextual knowledge and made available to AI systems for reasoning, decision-making, and action.

© Indium

13. What the Future Looks Like for Visionary Leaders

The future of Agentic Data Architecture will not look the same for every enterprise. The right path will depend largely on an organization's data maturity, existing technology landscape, and ability to adapt its architecture for AI-driven operations.

For large enterprises, the priority will be to build on existing investments rather than replacing them. Organizations with mature data platforms and modernization programs can incrementally introduce capabilities such as ontologies, semantic layers, knowledge graphs, and agentic services into their existing ecosystems. Their advantage lies in the ability to experiment with emerging technologies, validate their business value, and scale what works without disrupting established operations.

For mid-market enterprises, the opportunity is different. Organizations that are still early or midway through their modernization journey have greater flexibility to make foundational architecture decisions with AI-readiness in mind. Rather than retrofitting legacy environments later, they can select data platforms, governance models, semantic architectures, and knowledge frameworks that are designed to support Agentic AI from the outset.

Therefore, the direction of travel is clear: autonomous capabilities will increasingly become embedded in enterprise applications and business processes. The strategic question for leaders is not whether Agentic AI will reshape enterprise operations, but whether their organizations will have the architectural foundation required to benefit from it.

14. Conclusion

The evidence presented throughout this paper converges on a clear conclusion: enterprises fail to scale their AI models, either at the pilot stage or during enterprise-wide implementation, not because their AI models are inefficient, but because their enterprise data foundation lacks the intelligence required to make that data truly useful.

The path forward, however, will not necessarily be linear. Enterprises that are rapidly evolving from legacy architectures toward agentic systems capable of learning, adapting, deciding, and acting autonomously will be better positioned to thrive in the modern era of AI-powered transformation.

© Indium

Today, enterprises remain cautious about committing to long-term transformation programs as technologies, architectures, and operating models continue to mature. The objective, therefore, should not be to pursue autonomy for its own sake, but to build an architecture capable of progressively supporting greater autonomy without compromising trust, governance, or human judgment.

The enterprise of the future will not be defined simply by how many AI models it deploys, how many copilots it creates, or how many processes it automates. It will be defined by how effectively it transforms enterprise data into knowledge—and knowledge into action.

© Indium

References

[1] Gartner. “Gartner Predicts Over 40% of Agentic AI Projects Will Be Canceled by End of 2027.” Press release, June 25, 2025. https://www.gartner.com/en/newsroom/press-releases/2025-06-25-gartner-predicts-over-40-percent-of-agentic-ai-projects-will-be-canceled-by-end-of-2027

[2] Gartner. “Lack of AI-Ready Data Puts AI Projects at Risk” (interview with Roxane Edjlali; survey of 248 data management leaders, Q3 2024). Press release, February 26, 2025. https://www.gartner.com/en/newsroom/press-releases/2025-02-26-lack-of-ai-ready-data-puts-ai-projects-at-risk

[3] MIT Media Lab, Project NANDA. “The GenAI Divide: State of AI in Business 2025.” July 2025. (Findings: 95% of enterprise GenAI pilots show no measurable P&L impact; ~5% of custom tools reach production; US$30–40B enterprise investment.)

[4] Accenture Research. “AI-Ready Data for Advanced AI: From Capture to Value.” 2026. https://www.accenture.com/us-en/insights/ai-data/ai-ready-data

[5] S&P Global Market Intelligence. Enterprise AI adoption research, 2025. (Average organization scrapped 46% of AI proofs-of-concept before production.)

[6] Gartner. “Data Quality: Why It Matters and How to Achieve It.” (Poor data quality costs organizations an average of US$12.9 million per year.) https://www.gartner.com/en/data-analytics/topics/data-quality

[7] Sequeda, J., Allemang, D., & Jacob, B. “A Benchmark to Understand the Role of Knowledge Graphs on Large Language Models’ Accuracy for Question Answering on Enterprise SQL Databases.” data.world AI Lab / Proceedings of GRADES-NDA, 2024. https://arxiv.org/abs/2311.07509

© Indium

[8] Sequeda, J., & Allemang, D. “Increasing the LLM Accuracy for Question Answering: Ontologies to the Rescue!” data.world AI Lab, 2024. https://arxiv.org/abs/2405.11706

[9] Gartner. “Hype Cycle for Agentic AI, 2026.” (17% of organizations have deployed AI agents; 60%+ plan to deploy within two years.) https://www.gartner.com/en/articles/hype-cycle-for-agentic-ai

[10] Reuters (May 2025) and J.P. Morgan. “How AI Is Boosting Payments Efficiency and Cutting Fraud.” (≈US$1.5B in AI-attributed value; account-validation rejection rates reduced 15–20%.) https://www.jpmorgan.com/insights/payments/security-trust/ai-payments-efficiency-fraud-reduction

[11] Mastercard. “Mastercard Supercharges Consumer Protection with Gen AI” (Decision Intelligence Pro). Press release, February 2024. https://newsroom.mastercard.com/news/press/2024/february/mastercard-supercharges-consumer-protection-with-gen-ai/

[12] Adams, R., Henry, K.E., et al. “Prospective, Multi-Site Study of Patient Outcomes After Implementation of the TREWS Machine Learning-Based Early Warning System for Sepsis.” Nature Medicine 28 (2022): 1455–1460. https://www.nature.com/articles/s41591-022-01894-0

[13] McKinsey & Company. “The State of AI: Global Survey.” 2025. (88% of organizations use AI in at least one function; about one-third have begun scaling enterprise-wide.)

© Indium

INDIUM

About Indium

We are an AI services company specializing in Agentic AI, Data & Analytics, Application Engineering, and Quality Engineering.

By combining autonomous AI systems, modern engineering, and data intelligence, we build innovative AI solutions that drive measurable business outcomes and long-term value.

With 5,000+ associates worldwide, Indium partners with Fortune 500 companies, Global 2000 enterprises, and leading technology firms in Financial Services, Healthcare, Manufacturing, Retail, and Technology. Our teams operate in North America, India, the UK, Singapore, Australia, and Japan, helping businesses stay ahead in an AI-first world.

USA	INDIA	UK	SINGAPORE
Cupertino | Princeton Toll-free: +1-888-207-5969	Chennai | Bengaluru | Mumbai Hyderabad | Pune Toll-free: 1800-123-1191	London Ph: +44 1420 300014	Singapore Ph: +65 6812 7888

www.indium.tech

For Sales Inquiries
sales@indium.tech

For General Inquiries
info@indium.tech