dun & bradstreet

EBOOK

What Makes Data AI-Ready?

A Blueprint for Enterprises

Contents

INTRODUCTION

Laying a Trusted Foundation for Enterprise AI

AI has reached an inflection point. Its capabilities are now well established, and the challenge is no longer whether it works, but whether it can be trusted to operate at enterprise scale.

Many people who regularly use tools like ChatGPT, Gemini, or Perplexity have learned to expect occasional mistakes and hallucinations, and to recognize them quickly enough to work around them. That tolerance does not exist at enterprise scale. AI may be applied to high-stakes financial decisions, deployed within regulated industries, or asked to support work that spans borders and jurisdictions.

The risk grows further as AI becomes more autonomous. When AI moves beyond assisting decisions and begins to take action, errors do not remain isolated. Small inaccuracies can propagate through workflows, systems, and downstream decisions. The larger the organization, and the greater the autonomy of the AI, the more those errors can compound into material risk.

The primary obstacle to adoption of AI at enterprise scale is trust. The primary determinant of whether AI can be trusted is not the strength of the model, but the data that fuels it.

AI systems reflect the quality, structure, and reliability of the data they consume. Much of the data inside enterprises was not designed with AI in mind. Not all data is AI-ready.

This eBook addresses a simple but critical question:

What makes data AI-ready?

It examines the interdependent factors that determine whether data can support AI systems that are reliable, trustworthy, and therefore usable at enterprise scale.

01 **FOUNDATIONAL DATA INTEGRITY AND DETERMINISTIC GROUNDING:**

Data that is clean, deduplicated, consistently defined, and anchored to persistent identity.

02 **PROVENANCE, RIGHTS, AND SUPPLY CHAIN TRANSPARENCY:**

Clear lineage, ownership, and permitted use of data.

03 **STRUCTURAL AND INTERFACE READINESS FOR AI AND AGENTS:**

Data structured and exposed in ways that AI systems can reliably consume.

04 **AVAILABILITY AND INTEROPERABILITY ACROSS ENTERPRISE ENVIRONMENTS:**

Data accessible within and across the cloud and SaaS systems where AI operates.

05 **GOVERNANCE, COMPLIANCE, AND CONTROL:**

Embedded controls that support regulatory, contractual, and internal policy requirements.

01

Foundational Data Integrity and Deterministic Grounding

The first step in making data AI-ready is ensuring that the data itself is sound. At a basic level, this means the data accurately represents real-world entities and the relationships between them, and that those entities are grounded in persistent, deterministic identity.

In a probabilistic AI environment, grounding matters. Large language models infer and approximate. AI-ready data must provide stable identity anchors that prevent entities from being merged incorrectly, duplicated across systems, or fragmented over time. Persistent identifiers and deterministic entity resolution create that anchor.

Foundational data integrity also requires that records are clean, entities are deduplicated, relationships are correctly resolved, and definitions remain consistent over time. When these conditions are not met, AI systems operate on a distorted picture of reality.

Context and normalization are equally critical. Naming conventions, classifications, and attribute definitions must be harmonized so that different systems and agents interpret the same entity in the same way. Without consistent normalization, AI systems inherit ambiguity and produce inconsistent outputs.

Crucially, data integrity is not a one-time cleanup exercise. Enterprise data changes constantly as new records are added, entities evolve, and relationships shift. Maintaining integrity requires an ongoing data quality framework — one that continuously measures, reconciles, validates, and improves the dataset. AI-ready data depends on sustained operational discipline, not episodic remediation.

Foundational integrity and deterministic grounding are among the context engineering elements that provide a stable reference point that allows AI systems to reason consistently across use cases and over time. They are essential baseline conditions on which other dimensions depend.

85% of data leaders in 2026 say their organizations are adopting agentic AI, but 43% cite data readiness among their biggest obstacles.[1]

02

Provenance, Rights, and Supply Chain Transparency

Once data quality and identity grounding are established, the next requirement is provenance. Enterprises need to know where data originates, how it was collected, how it has been curated, and how it has moved through the data supply chain.

From provenance flows ownership, rights, and the ability to authorize specific uses of data in AI systems. Two organizations may hold data that appears identical in structure and quality. The difference lies in lineage and permitted use. One may rely on licensed data with limited visibility into its origins and constrained usage rights. Another may steward data end to end, with documented chain of custody and clearly defined entitlements.

AI-ready data must carry forward not only origin information, but also permitted use metadata. As AI systems combine and transform data across workflows, ambiguity around rights becomes operational risk. Clear entitlements enable authorized model training, enrichment, automation, and downstream AI applications.

Transparency and auditability reinforce this foundation. Enterprises must be able to trace how data was sourced, how it has been handled, and how it is being used by AI systems. This visibility supports compliance, internal policy enforcement, and accountability to regulators, partners, and customers.

In an AI context, data quality includes stewardship. Provenance anchors rights, governance, and trust in AI outputs, particularly as data is reused across systems and workflows.

03

Structural and Interface Readiness for AI and Agents

Structural readiness determines how AI systems access and interpret data. It does not determine whether the underlying data is accurate or authorized — that depends on the earlier dimensions. Instead, structural readiness ensures that AI systems can reliably consume and act on the data that has already been validated and governed.

For data to be AI-ready, it must be presented in a form that AI systems and autonomous agents can access and use consistently. This goes beyond basic availability through traditional APIs. Data must be structured and exposed in ways that preserve context, relationships, and operational constraints so that AI systems can interpret and act on it reliably.

Emerging protocols such as the Model Context Protocol (MCP) and agent-to-agent (A2A) standards define common interfaces between AI systems and enterprise data and tools. These protocols reduce the need for bespoke integrations and support scalability as AI use cases expand.

At the same time, structural readiness should not be confused with overall data readiness. Making data available through MCP or similar protocols signals that it is structured for AI use. Data quality, provenance, and authorization still determine whether that use is appropriate and trustworthy.

Recent McKinsey research found that 23% of enterprises are scaling an agentic AI system somewhere in their organization, and an additional 39% say they have begun experimenting with AI agents.[2]

04

Availability and Interoperability Across Enterprise Environments

Much of today's enterprise data lives inside cloud-based data platforms and SaaS applications. Many near-term AI use cases are embedded directly into these environments, supporting analytics, automation, and operational workflows.

Data cannot be considered AI-ready if it is not available within these environments in a practical, usable way. Availability in this sense is operational, not theoretical. Data may be well structured and properly governed, yet still require manual movement, custom pipelines, or one-off integrations to reach the systems where AI is being applied.

Interoperability addresses the reality that enterprise AI rarely operates within a single platform. Data must move between cloud data platforms, analytics environments, and enterprise applications without losing context, permissions, or governance. When data remains siloed or constrained by incompatible formats, AI usage becomes fragmented and difficult to scale.

AI readiness depends on whether data can be activated consistently wherever AI is embedded into day-to-day work.

A PwC survey found that 60% of top-performing organizations have broadly adopted cloud specifically for AI, compared with 27% of other companies.[3]

05

Governance, Compliance, and Control

As AI becomes embedded in enterprise workflows, governance becomes an operational requirement rather than a policy afterthought. Enterprises are accountable for how AI systems behave, what data they use, and how decisions are made.

Governance builds on identity, provenance, and entitlements. AI-ready data must support regulatory compliance, internal controls, and responsible use as part of normal operation.

In practice, this means enterprises need visibility and control over how data is accessed and applied. Lineage tracking, access management, usage constraints, and auditability make it possible to understand where data came from, how it is being used, and whether that use aligns with policy and regulation.

Data providers can support this work by embedding governance features into their data, but they do not replace enterprise responsibility. When governance is embedded at the data layer and carried into AI workflows, organizations can deploy AI more broadly with confidence. When governance is added after the fact, adoption slows under the weight of review and risk mitigation.

A 2025 Survey of UK enterprises showed that only 7% of organizations have fully embedded AI governance, despite 93% using AI in some capacity.[4]

IMPLICATIONS

What This Means for Different Stakeholders (Including Agents)

The five dimensions above define what makes data AI-ready.

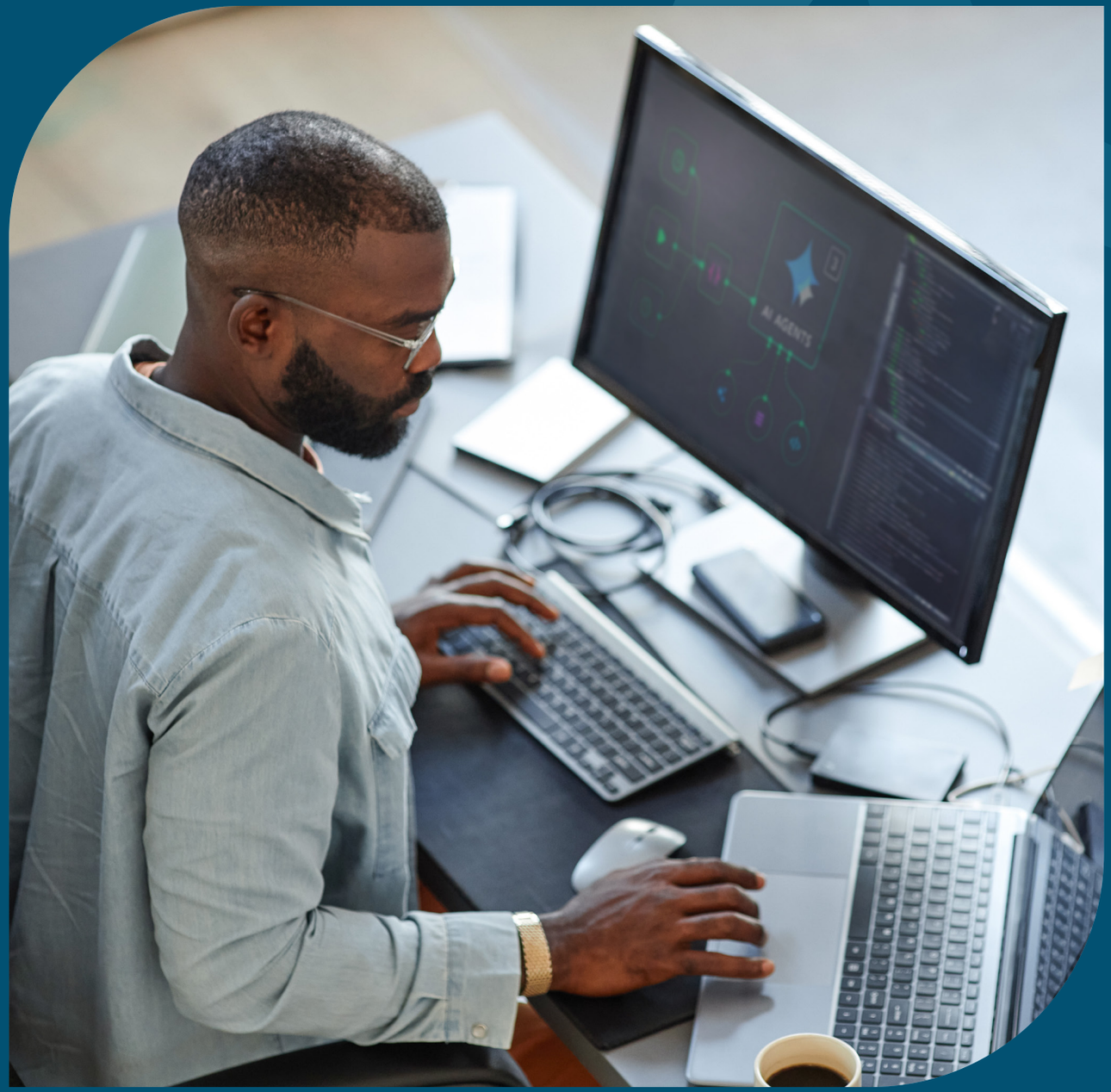
Their implications differ slightly depending on who is building, deploying, or relying on AI systems.

Enterprises

For enterprises, the priority is confidence at scale. AI must operate against a verified and governed representation of commercial reality, particularly as use cases move into credit, compliance, procurement, and other high-stakes workflows. Deterministic identity, clear provenance, structural readiness, interoperability, and embedded governance together determine whether AI can be trusted beyond experimentation.

Developers

For developers and AI builders, these same dimensions translate into practical requirements: stable entity resolution, consistent schemas, clear entitlements, and predictable access patterns. As AI systems and agents interact across platforms, they depend on persistent identifiers and normalized context to prevent entity drift, duplication, and unintended merges. AI-ready data reduces ambiguity at the system level, not just the user interface.



Agents

AI-readiness of data is different still for the agents themselves. Agents do not independently determine which data is authoritative; they are routed to sources through tool registration, retrieval layers, and enterprise governance controls. Data that is accurate but not exposed in a structured, machine-readable, and policy-aware way is effectively invisible to the agent. AI-ready data must therefore be not only verified and governed, but discoverable and selectable within the routing logic that determines how agents retrieve and act on information.

CONCLUSION

AI Adoption Starts at the Foundation

There is intense pressure on enterprises to move quickly with AI. Large investments have already been made, expectations for returns are rising, and the pace of innovation continues to accelerate. It is tempting to treat progress along the AI maturity curve as a race to adopt faster models or more advanced capabilities.

In practice, durable progress follows a different logic. A useful maxim captures it well: slow is smooth, and smooth is fast. Careful, disciplined execution at the foundation reduces friction later.

When identity is grounded, context is normalized, rights are clear, data is accessible, and governance is embedded, AI systems behave more predictably and scale more smoothly. When those foundations are weak, organizations spend time managing errors, exceptions, and risk, which ultimately slows progress.

AI adoption starts at the foundation. Once that foundation is established, it must be maintained and protected. Making data AI-ready is not the slow path. It is the path that allows enterprises to move faster with confidence — and to convert experimentation into sustained value.

---

[1] **Source**: Precisely, Fourth Annual Study Finds AI Confidence Outpaces Readiness as Data Integrity Gaps Persist, January 2026. Percentages reflect survey responses from data leaders regarding agentic AI adoption and data readiness challenges.

[2] **Source:** McKinsey & Company, The State of AI, November 2025. Percentages reflect survey responses from enterprises regarding the scaling of agentic AI systems at 23% and experimentation with AI agents at 39%. SOURCE

[3] **Source:** PwC, Cloud and AI Business Survey, 2024. Percentages reflect survey responses comparing top-performing organizations broadly adopting cloud for AI at 60% with other companies at 27%. SOURCE

[4] **Source:** Trustmarque, AI Governance Report, July 2025. Percentages reflect survey responses from UK enterprises regarding AI usage at 93% and the extent of fully embedded AI governance at 7%. SOURCE

ABOUT DUN & BRADSTREET®

Dun & Bradstreet provides the verified commercial identity foundation for enterprises to deploy AI at scale. The company originated the D-U-N-S® Number in 1963, now the global standard for identifying commercial entities. Anchored by this identifier, the D&B Commercial Graph™ structures and connects business identity consistently across systems, enabling AI to operate on accurate, validated data. Since 1841, businesses of every size have relied on Dun & Bradstreet to navigate change and accelerate growth. For more information, visit www.dnb.com.

dnb.com