

Making Your Data AI-Ready Through Data Products

How Data Products Give AI Leaders
an Unfair Advantage



Executive Summary

AI is moving faster than we ever imagined.

The companies that master AI-ready data will dominate the future, while those that don't will struggle to compete in an increasingly intelligent world. If your data isn't managed, governed, and productized for AI, you're already falling behind.

Data products offer a clear path forward.

AI is only as powerful as the data on which it is built. Yet many organisations struggle to ready their data for AI, resulting in long implementation cycles, unreliable models and governance risks. The key to AI success is to move away from fragmented, ad hoc data pipelines and embrace data products – they are structured, reusable, high-quality and enable AI-driven innovation.

A data product is more than just a dataset; it is a carefully designed, governed and maintained asset that ensures AI teams have access to trusted, compliant and scalable data. With the explosion of AI-driven use cases across industries, organisations urgently need to rethink their data strategy or risk being left behind.

This whitepaper breaks down the biggest obstacles to AI readiness and explains how data products solve them. You'll learn why AI needs structured, governed data, how data products create a seamless flow between AI and data, and how to implement a practical, AI-ready data product strategy. By the end, you'll have a clear roadmap to speed up AI deployment, ensure compliance, and scale AI with confidence – and with the right data.

The message is clear: The companies that embrace data products today will dominate the AI-driven economy of tomorrow. The ones that don't will struggle to compete.

"AI-ready data means that your data must be representative of the use case, of every pattern, errors, outliers and unexpected emergence that is needed to train or run the AI model for the specific use."

Gartner®

1

The AI Readiness Problem: Why Most AI Initiatives Struggle



Despite significant investments in AI, many organizations face an uphill battle when it comes to efficiently deploying AI solutions. They quickly hit a fundamental roadblock: their data is not AI-ready. What does this mean in practice?

It means that instead of AI being a driver of efficiency and transformation, teams are bogged down by poor data quality, governance gaps, security risks and compliance challenges.

"Through 2026, 30% of generative AI (GenAI) projects will be abandoned after proof of concept due to poor data quality, inadequate risk controls, escalating costs or unclear business value."

Gartner®

The Hidden Cost of Poor Data Quality

Data fuels AI, but just as poor quality fuel can damage an engine, poor quality data can render AI models ineffective or even harmful. Data that is inconsistent, incomplete or duplicated introduces inefficiencies and inaccuracies into AI models. Skewed or biased data sets lead to AI decisions that lack fairness, trust and reliability. Without standardised and structured data, AI teams are forced to spend months wrestling with data - time that should be spent developing models and innovating.

According to a **recent study by Gartner®**, poor data quality costs organizations an average of \$12.9 million per year. These costs extend beyond wasted effort—poor-quality data leads to failed AI projects, lost revenue opportunities, and reputational damage from biased or unreliable AI outputs.

"Only 10-30% of most data in enterprises is structured. The rest is unstructured."

Gartner® →

Governance Gaps and the Challenge of Trust

For AI to succeed at scale, organizations need complete visibility into the lineage, quality, and ownership of their data. But in most organizations, data is spread across silos with no clear governance model, accountability, or tracking. AI teams often struggle to determine whether they can trust a data set, let alone use it in production. Who owns the data? Where did it come from? Has it been modified in a way that could pose a risk? Without governance mechanisms in place, AI development slows to a crawl.

70% of AI high performers have experienced difficulties with data, including challenges in defining processes for data governance.

McKinsey® →

We all see how AI regulation is evolving rapidly (e.g. with the EU AI Act). Organizations that fail to prioritize governance will soon find themselves unable to deploy AI solutions in compliant and ethical ways.

Security, Compliance, and Ethical Considerations

AI applications increasingly rely on sensitive customer data, proprietary business information, and personally identifiable information (PII).

This presents significant risks if organizations do not enforce strong security, access control, and compliance measures. Without role-based access control, encryption, and auditability, AI models could inadvertently expose sensitive information, resulting in massive regulatory fines or breaches of customer trust.

AI is also under increasing scrutiny for its ethical implications. Models trained on unregulated, unmonitored data sets can amplify biases, make unfair predictions, or even violate emerging AI regulations such as the EU AI Act. Organizations that treat data governance as an afterthought will struggle to keep pace with regulatory expectations, and risk having their AI initiatives derailed by compliance issues.

The AI Deployment Bottleneck

AI should be moving at the speed of innovation, but for most organizations, AI deployment is slow, complex, and expensive.

Data teams spend an estimated 80% of their time on data preparation, rather than actual AI development. Each new AI use case requires a bespoke data integration effort, further increasing the time and cost to deploy new AI models.

Meanwhile, as AI adoption accelerates, organizations that fail to streamline data preparation will fall further behind competitors who can deploy AI models faster and at lower cost. The writing is on the wall: without AI-ready data, AI innovation will remain stuck in first gear.

2

The Solution: Data Products as the Foundation for AI-Ready Data

The key to overcoming these challenges is to move away from ad hoc data pipelines and treat data as a product. This shift fundamentally changes the way organizations work with data – ensuring that data is high quality, governed, secure, and optimized for AI consumption from day one.



What Makes a Data Product?

A data product is a curated, reusable dataset that is:

- Structured & Standardized: AI-ready by design, eliminating the need for additional processing.
- Governed & Secure: Ensures compliance, ownership, and access control.
- Discoverable & Reusable: Eliminates duplication and fosters data sharing across teams.
- Optimized for AI: Enriched with metadata, monitoring, and versioning to ensure long-term reliability.

McKinsey → further emphasizes the significance of data products in scaling AI initiatives:

“Data products are the secret sauce for scaling AI. They help deliver data-intensive applications as much as 90 percent faster, at 30 percent lower cost, and with a reduced risk and data governance burden.”

This underscores that organizations that implement data products can achieve significant improvements in speed, cost efficiency, and risk management, thereby improving their AI capabilities and overall performance.

By embedding quality, governance, and security into every data product, organizations remove the barriers that slow AI and set the stage for scalable, high-impact AI deployments.

3

The Data-AI Flywheel: How AI and Data Products Reinforce Each Other

AI and data products are not only connected, they feed each other in a continuous innovation loop. AI needs high-quality data to work effectively, but AI can also be used to improve and maintain data products over time.



How AI Enhances Data Products

- Automated Data Cleaning & Labeling: AI can detect anomalies, correct errors, and enrich datasets without manual intervention.
- Intelligent Metadata Generation: AI can automatically generate metadata that improves discoverability and usability of data products.
- Bias Detection & Fairness Monitoring: AI-driven tools can continuously assess datasets for bias, improving model fairness.

How Data Products Improve AI

- Higher Accuracy & Reliability: AI models trained on well-structured data products deliver more accurate and explainable results.
- Scalability Across Use Cases: Standardized data products allow organizations to build AI applications faster without starting from scratch each time.
- Better AI Governance & Compliance: Clear lineage and provenance ensure that AI models meet regulatory and ethical standards.

This data-AI flywheel is the key to long-term AI scalability. Organizations that embrace this model will have a sustainable, continuously improving AI ecosystem, while those that don't will struggle to compete in an AI-driven world.

Conclusion:

The Future of AI Belongs to Those Who Prioritize Data Products

The companies that take proactive steps to make their data AI-ready will shape the future. Those that ignore the urgency of structured, governed, and scalable data will be playing catch-up. AI is here. Now is the time to turn data into a productive asset that drives AI success.