LUMO INSIGHTS

Thursday, 16 July 2026

# Data strategy: The deciding factor in AI success

## Executive Summary

In the past week, a series of developments underscored that the biggest barrier to AI success is not the algorithms - it's the data. From new regulations tightening control of training data to fresh evidence of poor data quality derailing AI projects, the message is clear: enterprises that get their data and architecture right are pulling ahead, while others stumble. This briefing distills the latest insights on how data strategy and infrastructure are determining who wins with AI - and what C-suite leaders should do next.

## Leaders vs. laggards: Data strategy sets the pace

A widening performance gap in AI adoption is becoming evident between organisations that treat data as a strategic asset and those that do not. A new industry report suggests the next phase of enterprise AI competition may depend less on who has the most advanced models and more on how businesses manage and integrate their data - especially the troves of unstructured information like documents, emails, and media files (www.unite.ai [1]). Early AI adopters ("future-built" leaders) are outpacing laggards in measurable results, in part because they invested in stronger data foundations. A Boston Consulting Group study found that these AI leaders expect double the revenue growth and 40% greater cost reductions from their AI initiatives compared to those trailing behind (www.forbes.com [2]). In short, the companies winning with AI began by investing in better data architecture and governance, not just better algorithms.

This reality was highlighted in an announcement by ServiceNow, where a top executive pointed out that most enterprise AI failures aren't due to subpar models at all - they fail because data remains trapped in silos and is poorly governed right where AI needs it most (secure.businesswire.com [3]). When customer data, operations data, and other critical information are fragmented across disconnected systems, AI can only offer "shallow" insights that suggest actions but can't reliably execute them (secure.businesswire.com [4]). This fragmentation creates blind spots and erodes trust in AI outputs.

The organisations pulling ahead are tackling this problem by breaking down data silos. They are unifying data across cloud and on-premises systems and feeding their AI models with real-time, context-rich information drawn from every corner of the business. As one industry leader put it, enterprises that win the AI race bring "trusted, contextual data directly into the workflows that run the business," giving both teams and algorithms the insight to act with confidence (secure.businesswire.com [5]). In practice, that means modernising data architecture - for example, building unified data platforms (like the emerging "lakehouse" systems) where a single copy of data can serve both analytics and operational needs in real time (www.databricks.com [6]). By investing in such integrated data environments under a common governance framework, companies ensure their AI systems are never starved for high-quality information. The payoff is tangible: faster decisions, more automated processes, and AI-driven innovations that actually stick, rather than stalling out after

impressive pilots.

References:

[1] www.unite.ai — https://www.unite.ai/nasunis-the-state-of-enterprise-file-data-annual-report-2026-finds-enterprise-ai-adoption-is-outpacing-data-readiness/#:~:text=decision,separated%20from%20broader%20data%20management
[2] www.forbes.com — https://www.forbes.com/sites/sylvainduranton/2025/11/20/the-widening-ai-value-gap-leaders-profit-laggards-get-left-behind/#:~:text=stark%20story.%20Future,plus%20sectors.%20They
[3] secure.businesswire.com — https://secure.businesswire.com/news/home/20260506850477/en/ServiceNow-launches-the-real-time-data-foundation-that-puts-autonomous-AI-to-work-across-the-enterprise#:~:text=agent%20governance%20that%20autonomous%20AI,party
[4] secure.businesswire.com — https://secure.businesswire.com/news/home/20260506850477/en/ServiceNow-launches-the-real-time-data-foundation-that-puts-autonomous-AI-to-work-across-the-enterprise#:~:text=agent%20governance%20that%20autonomous%20AI,party
[5] secure.businesswire.com — https://secure.businesswire.com/news/home/20260506850477/en/ServiceNow-launches-the-real-time-data-foundation-that-puts-autonomous-AI-to-work-across-the-enterprise#:~:text=with%20every%20workflow%2C%20making%20AI,vice%20president%20and%20general%20manager
[6] www.databricks.com — https://www.databricks.com/company/newsroom/press-releases/databricks-launches-ltap-first-lake-transactionalanalytical#:~:text=Databricks%20today%20launched%20LTAP%20,Data%20and%20AI%20company%2C%20today

## The unstructured Data bottleneck

Even as organisations pour resources into AI projects, many are discovering that years of neglected data issues are catching up with them. A new survey of 1,000 IT and procurement leaders revealed that while 97% of large firms have already started deploying or piloting AI agents, over half of those projects are failing to meet their goals due to unmanageable unstructured data (techedgeai.com [1]). Unstructured data - the documents, emails, images, logs, and other information that don't neatly fit into databases - has exploded in volume, yet only 16% of companies currently treat managing this data as a strategic IT priority. This oversight has led to a situation where AI models trained in "lab" conditions falter in the real world, encountering messy, fragmented data environments they weren't prepared for. One data executive dubbed this the "AI ROI cliff," where promising AI prototypes crash in production because they can't scale a mountain of poor-quality data and siloed systems.

The consequences of this data readiness shortfall are showing up in hard metrics. Gartner analysts have predicted that through this year, 60% of AI initiatives lacking "AI-ready" data will ultimately be abandoned mid-project, and we're already seeing this play out - 42% of U.S. enterprises report they have had to scrap at least one AI project due to data problems (www.beri.net [2]). Likewise, Nasuni's 2026 State of Enterprise File Data report found only 43% of AI initiatives fully achieve their intended objectives, with the rest underdelivering or failing outright, largely because of data issues (techedgeai.com [3]). The same study found 90% of organisations cite challenges with data quality, integration, security, or trust as major barriers to scaling AI beyond early pilot projects (techedgeai.com [4]). When critical business information is incomplete, inconsistent, or locked away in incompatible systems, AI outcomes become unreliable - and stakeholders lose faith.

The good news is that awareness of this bottleneck is prompting action. Many enterprises are now urgently investing in modern data management and integration solutions to tame their unstructured data sprawl. According to the Nasuni survey, 60% of companies plan to increase spending on unstructured data management in the next 18 months (www.unite.ai [5]), reflecting a recognition that improving data organisation, enrichment, and accessibility is essential for AI success. Business leaders are also prioritising data quality initiatives - cleaning up duplicates, standardising formats, and fortifying data governance processes - to prevent "garbage in, garbage out" scenarios that doom AI projects. The takeaway for the C-suite is clear: without clean, well-governed data that's easily accessible to AI systems, even the most powerful algorithms will stumble.

References:

[1] techedgeai.com — https://techedgeai.com/nasunis-2026-state-of-enterprise-file-data-report-shows-unstructured-data-as-the-biggest-ai-bottleneck/#:~:text=Blocker%2C%20Driving%20a%20Surge%20in,sheer%20volume%20and%20complexity%20of
[2] www.beri.net — https://www.beri.net/article/ai-project-failure-complete-

guide-2026#:~:text=Gartner%27s%20prediction%3A%2060,project%20due%20to%20data%20issues
[3] techedgeai.com — https://techedgeai.com/nasunis-2026-state-of-enterprise-file-data-report-shows-unstructured-data-as-the-biggest-ai-bottleneck/#:~:text=drift%20and%20compliance%20risk,Nasuni%E2%80%99s%20patented%20architecture%E2%80%94combining%20a
[4] techedgeai.com — https://techedgeai.com/nasunis-2026-state-of-enterprise-file-data-report-shows-unstructured-data-as-the-biggest-ai-bottleneck/#:~:text=drift%20and%20compliance%20risk,Nasuni%E2%80%99s%20patented%20architecture%E2%80%94combining%20a
[5] www.unite.ai — https://www.unite.ai/nasunis-the-state-of-enterprise-file-data-annual-report-2026-finds-enterprise-ai-adoption-is-outpacing-data-readiness/#:~:text=unstructured%20data%20management%20all%20emerged,The

## Data as IP: protection and proprietary advantage

As AI becomes embedded in products and decisions, the data feeding these models is turning into precious intellectual property. Companies increasingly view their stockpiles of proprietary data - from customer interactions to supply chain metrics - as unique strategic assets and competitive moats that must be safeguarded. In fact, tech leaders are openly voicing concerns about protecting enterprise data from being inadvertently shared or commoditised in the AI era. To this point, Microsoft's new "Frontier" AI initiative, announced this month, puts heavy emphasis on preserving customers' data and knowledge as exclusive intellectual property. The company has pledged that a client's data, models, and know-how will not be used to train broader AI systems in ways that could erode that client's competitive advantage (blogs.microsoft.com [1]). Microsoft CEO Satya Nadella was even more direct, saying there is *"no societal permission for an AI future that eats the intelligence of the companies it's deployed inside"* (blogs.microsoft.com [2]). In other words, organisations will only embrace AI if they're confident it won't "steal" their secret sauce. This stance reflects a broader industry shift towards treating data, and the insights derived from it, as sacred capital.

Data ownership questions are no longer just IT issues - they are now boardroom issues. Real-world examples continue to show that failing to clarify who owns and controls data can sink AI initiatives. A telling anecdote circulating in IT leadership circles describes an AI pilot that dazzled in testing but later died in committee because stakeholders couldn't agree on data ownership and usage rights. Forward-thinking companies are avoiding such debacles by establishing clear data governance frameworks upfront - defining who is accountable for data quality, access permissions, and ethical use. They recognise that every new AI application raises questions about data provenance and responsibility: Who collected the data? Is it compliant with privacy laws and customer expectations? Who can use or share it? By proactively answering these questions, companies create a solid legal and ethical footing for AI projects and reduce the risk of later complications.

References:

[1] blogs.microsoft.com — https://blogs.microsoft.com/blog/2026/07/02/microsoft-frontier-company-ai-engineering-that-amplifies-and-protects-your-intelligence/#:~:text=our%20Global%20SI%20partners%2C%20including,diverse%2C%20open%2C%20heterogeneous
[2] blogs.microsoft.com — https://blogs.microsoft.com/blog/2026/07/02/microsoft-frontier-company-ai-engineering-that-amplifies-and-protects-your-intelligence/#:~:text=customer%E2%80%99s%20IQ%20is%20protected,Microsoft%E2%80%99s%20platform

## Regulatory wake-up call: compliance driving Data discipline

In the past week, regulators sent a strong signal that organisations must get their data houses in order if they want to innovate with AI. On July 8, Europe's top data protection authority issued new guidelines making it explicit that the EU's GDPR privacy law applies fully to AI training data scraped from the web (www.techtimes.com [1]). This ruling closes a loophole that many AI developers had informally relied on - using publicly available internet data for training without explicit consent. Now, any personal data for EU citizens included in training sets must have a legal basis and meet strict standards like purpose limitation and anonymity. For global businesses, this means AI data practices need immediate review. Non-compliance could invite regulatory action or hefty fines, as European officials zero in on how AI companies handle personal information.

This European stance is part of a larger trend of rising scrutiny on data in AI. Around the world, governments and standards bodies are debating how to enforce transparency, fairness, and sovereignty over data used in AI systems. Companies are having to navigate a patchwork of regulations - from local data residency laws to sector-specific rules on sensitive data - all while trying to harness data for competitive advantage. The push for “data sovereignty” is evident in moves like new cloud certification programmes in Europe aimed at keeping sensitive data within national borders (www.cio.com [2]). The message to enterprises is that embracing AI goes hand-in-hand with strengthening data governance. By investing in compliance-ready data architectures (think robust data catalogues, audit trails, and permission controls) and adopting privacy-preserving techniques, organisations can both satisfy regulators and build greater trust with customers. In turn, this trust becomes an enabler for AI: users and regulators alike are far more comfortable with AI innovations when they know the underlying data is handled responsibly.

References:
[1] www.techtimes.com — https://www.techtimes.com/articles/320024/20260709/gdpr-applies-ai-training-data-eu-ends-web-scraping-free-pass-every-lab.htm#:~:text=the%20most%20contested%20legal%20questions,EU%20framework%20to%20address
[2] www.cio.com — https://www.cio.com/data-governance/#:~:text=governance%20playbook%20obsolete%20By%20Volker,Management%20Opinion%20Shadow%20agents%3A%20How

## Key Statistics

- Over 50% - More than half of large-company AI projects are falling short of their goals, despite 97% of big firms piloting AI agents, due to unmanaged, fragmented unstructured data ([techedgeai.com](https://techedgeai.com/nasunis-2026-state-of-enterprise-file-data-report-shows-unstructured-data-as-the-biggest-ai-bottleneck/#:~:text=Blocker%2C%20Driving%20a%20Surge%20in,sheer%20volume%20and%20complexity%20of)).
- 42% - In a recent survey, 42% of U.S. enterprises said they had abandoned at least one AI project because of data quality or availability issues ([www.beri.net](https://www.beri.net/article/ai-project-failure-complete-guide-2026#:~:text=Gartner%27s%20prediction%3A%2060,project%20due%20to%20data%20issues)).
- 90% & 43% - Ninety per cent of companies cite data integration, security and trust challenges as a barrier to AI scale, and only 43% of AI initiatives fully achieve their intended outcomes ([techedgeai.com](https://techedgeai.com/nasunis-2026-state-of-enterprise-file-data-report-shows-unstructured-data-as-the-biggest-ai-bottleneck/#:~:text=drift%20and%20compliance%20risk,Nasuni%E2%80%99s%20patented%20architecture%E2%80%94combining%20a)).
- 80% - A study of 500+ firms using AI agents found 80% already see measurable ROI, with those that treat agents as core infrastructure (not experiments) reaping the greatest benefits ([claude.com](https://claude.com/blog/how-enterprises-are-building-ai-agents-in-2026#:~:text=for%20research%20and%20reporting%20over,analysis%20from%205%20hours%20to)).

KEY TAKEAWAY

AI success now depends less on model breakthroughs and more on robust data foundations. Those leading the pack are investing in integrated, well-governed, real-time data architectures and protecting their proprietary data - a strategy that drives ROI and builds a lasting competitive edge.

## Sources


GDPR Applies to AI Training Data: EU Ends Web Scraping Free Pass for Every Lab
https://www.techtimes.com/articles/320024/20260709/gdpr-applies-ai-training-data-eu-ends-web-scraping-free-pass-every-lab.htm

Nasuni's The State of Enterprise File Data Annual Report 2026 finds enterprise AI adoption is outpacing data readiness
https://www.unite.ai/nasunis-the-state-of-enterprise-file-data-annual-report-2026-finds-enterprise-ai-adoption-is-outpacing-data-readiness/

AI project management: Why traditional methods fall short
https://www.informationweek.com/machine-learning-ai/why-traditional-project-management-doesn-t-work-for-ai-projects

ServiceNow launches the real-time data foundation that puts autonomous AI to work across the enterprise
https://www.businesswire.com/news/home/20260506850477/en/ServiceNow-launches-the-real-time-data-foundation-that-puts-autonomous-AI-to-work-across-the-enterprise

Microsoft Frontier Company: AI engineering that amplifies and protects your intelligence
https://blogs.microsoft.com/blog/2026/07/02/microsoft-frontier-company-ai-engineering-that-amplifies-and-protects-your-intelligence/

How enterprises are building AI agents in 2026
https://claude.com/blog/how-enterprises-are-building-ai-agents-in-2026

The Widening AI Value Gap: Leaders Profit, Laggards Get Left Behind
https://www.forbes.com/sites/sylvainduranton/2025/11/20/the-widening-ai-value-gap-leaders-profit-laggards-get-left-behind/