LUMO INSIGHTS

Thursday, 9 July 2026

Data, not models, is now the deciding factor in AI success

Executive Summary

A wave of new research, real-world examples, and product releases in recent days underscores a common theme: the key to AI competitiveness now lies in data strategy and architecture. Companies that invest in robust, well-governed data foundations are reaping significantly higher returns on AI, while those that neglect data quality, integration, and ownership are falling behind - and facing increasing regulatory and operational risks.

Data Infrastructure: The new AI divide

Multiple reports and industry moves this week point to a growing chasm between AI leaders and laggards - and the divide isn't about algorithms, but data infrastructure. A new global survey by Confluent finds that nearly three-quarters (72%) of IT executives cite inadequate real-time data infrastructure as the top factor holding back their AI initiatives (www.confluent.io [1]). In other words, the most advanced AI model won't deliver ROI if it's starved of timely, well-integrated data. This resonates with a broader industry realisation that AI's limitations today are less about model innovation and more about the data pipelines feeding those models.

Supporting this view, a CData "State of AI Data Connectivity" study revealed that only 6% of enterprise AI leaders consider their data infrastructure fully ready for AI - highlighting a massive readiness gap (www.cdata.com [2]). The same study draws a direct link between data infrastructure maturity and AI success: every high-maturity AI organisation had built a centralised, semantically consistent data integration layer, while 80% of low-maturity companies hadn't even started on this foundation (www.cdata.com [3]). The lesson is clear for business leaders - robust data architecture is emerging as the decisive factor in who derives value from AI at scale.

Even the focus of AI investment is shifting from algorithms to data plumbing. According to the CData report, only 9% of organisations now rank developing AI models as their top priority, whereas 83% are investing in unified, real-time data access layers to support AI (www.cdata.com [4]). Similarly, Box's just-released 2026 State of AI in the Enterprise survey found that the key differentiators between AI frontrunners and stragglers are "content access, governance, and platform flexibility" - not superior models (venturebeat.com [5]). Together, these findings underscore that AI leadership in 2026 is defined by having the right data, in the right place, at the right time, wrapped in the right architecture.

References:

[1] www.confluent.io — https://www.confluent.io/press-release/2026-data-streaming-report/#:~:text=Calif.,examines%20the%20challenges%20that%20enterprises

[2] www.cdata.com — https://www.cdata.com/company/press/state-of-ai-data-connectivity-report/#:~:text=Only%206,which%20draws%20on%20independently%20collected

[3] www.cdata.com — https://www.cdata.com/company/press/state-of-ai-data-connectivity-report/#:~:text=Infrastructure%20maturity%20is%20the%20great,fundamental%20shift%20in%20AI%20strategy

[4] www.cdata.com — https://www.cdata.com/company/press/state-of-ai-data-connectivity-report/

#:~:text=Only%209,It%27s%20determined%20by%20the
[5] venturebeat.com — https://venturebeat.com/orchestration/box-survey-why-enterprise-ai-leaders-are-outperforming-their-peers#:~:text=Content%20access%2C%20governance%2C%20and%20platform,a%20notable%20return%20on%20their

Data quality & governance - from bottleneck to boardroom

Data quality and governance challenges are proving to be the Achilles' heel of many AI projects. Industry analyses continue to estimate that as many as 70 - 85% of AI initiatives fail to deliver their intended value (plusoperator.com [1]). In the realm of generative AI, for instance, over half of projects were abandoned post-pilot by late 2025 - often killed by poor data quality, lack of risk controls, or ballooning costs. These sobering numbers reinforce a hard truth: if your data is unreliable, fragmented, or ungoverned, even state-of-the-art AI will stumble.

A real-world example hit headlines recently when Starbucks was forced to scrap an AI-powered inventory management system after just nine months, due to consistent errors and integration woes. The automated system - intended to optimise restocking across 11,000 stores - ended up miscounting products (confusing items like similar milk varieties) and couldn't account for local factors like weather or special events (finance.yahoo.com [2]). Store managers lost trust in the AI as it overrode their local expertise and didn't play nicely with existing point-of-sale systems. The project was quietly shelved, illustrating how insufficient data accuracy and poor systems integration can turn a high-profile AI investment into wasted effort.

These failures are fuelling a new urgency around data governance, quality control, and clear data ownership. Confluent's study reported that 66% of IT leaders struggle with uncertainty about data lineage, timeliness, and quality in their organisations (www.confluent.io [3]) - essentially, they lack full trust in their data. Two-thirds also pointed to fragmented data ownership as a major obstacle, highlighting how unclear accountability for data can stall AI progress (www.confluent.io [4]). The message for the C-suite is that ensuring clean, well-understood data isn't just an IT problem, but a strategic imperative to achieve AI-driven outcomes.

Regulators are also raising the stakes. Europe's AI Act, set to begin enforcement in August 2026, will impose strict transparency and data management requirements on high-risk AI systems - with fines up to €35 million or 7% of global revenue for non-compliance (www.legiscope.com [5]). Meanwhile, continued GDPR enforcement and new data sovereignty laws worldwide (from data localisation mandates to stricter privacy rules) mean that companies must know exactly where their data resides and how it's used (lyceum.technology [6]). In this environment, strong data governance is no longer optional; it's now a boardroom-level concern, vital for both unlocking AI value and avoiding legal pitfalls.

References:

[1] plusoperator.com — https://plusoperator.com/around-85-of-data-science-projects-fail-how-not-to-become-a-number-in-this-statistic/#:~:text=Why%2085,never%20make%20it%20to%20production
[2] finance.yahoo.com — https://finance.yahoo.com/markets/stocks/articles/reuters-starbucks-ends-ai-inventory-135116717.html#:~:text=reported%20in%20February%20that%20the,and%20other%20beverage%20products%2C%20was
[3] www.confluent.io — https://www.confluent.io/press-release/2026-data-streaming-report/#:~:text=in%20AI%20itself,cite%20data%20infrastructure%20and%20data
[4] www.confluent.io — https://www.confluent.io/press-release/2026-data-streaming-report/#:~:text=insufficient%20infrastructure%20for%20real,cite%20data%20infrastructure%20and%20data
[5] www.legiscope.com — https://www.legiscope.com/blog/eu-ai-act-timeline-deadlines.html#:~:text=human%20oversight,obligations%20%2B%20national%20authority%20designation
[6] lyceum.technology — https://lyceum.technology/magazine/data-sovereignty-requirements-ai-by-country/#:~:text=Data%20Sovereignty%20Requirements%20for%20AI,2026%20introduced%20aggressive%20new%20frameworks

New platforms and the Data moat

As AI models become commoditised, enterprises are increasingly looking to data itself as the source of competitive advantage. In fact, Gartner now characterises cutting-edge AI models as "strategic commodities," meaning gains from model superiority alone are fleeting (aiireland.ie [1]). The only truly defensible assets, and thus the new “moats” in the AI era, are proprietary datasets and the robust pipelines to use them. Companies that can gather unique, high-quality data - and weave it into their AI - will enjoy an edge that rivals can’t easily copy.

Forward-looking organisations are moving aggressively on this front. Many chief data officers (CDOs) are spearheading efforts to catalogue and clean internal data, break down silos, and establish clear data ownership structures. As one analysis put it, failing to govern, protect, and leverage your company’s unique data is tantamount to handing competitors your advantage on a silver platter (aiireland.ie [2]). This mindset shift is evident in the surge of interest around concepts like *data mesh* and *data fabric*, which treat data as a product and emphasise cross-functional ownership and standardisation. The goal is a living, well-documented data estate that any AI system - present or future - can reliably tap into.

The tech industry is responding with tools to help enterprises build these data foundations. This week saw announcements aimed at fusing once-disparate data capabilities into unified AI-ready platforms. For example, Couchbase launched a new **AI Data Plane** - a single data layer spanning from cloud to edge - to give AI "agent" systems a governed, real-time memory and context source for their decisions (www.storagenewsletter.com [3]). Similarly, Databricks used its recent Data + AI Summit to recast its Lakehouse architecture as an **AI control plane** for the enterprise, where autonomous agents can securely access and act on data across the organisation (inferenz.ai [4]). Even cloud giants are in the mix: Microsoft’s Build 2026 announcements included a vector-optimised database service and a feature called Fabric IQ to help define business-specific knowledge for AI agents (www.digitaltoday.co.kr [5]). All these developments reflect a common strategic direction - making data accessibility, quality, and governance a solved problem so that AI initiatives can move from sandbox to at-scale production.

For senior technology leaders, the takeaway is clear. The winners in the AI race are those treating data as a first-class strategic asset: investing in modern data architectures like lakehouses or data meshes, integrating real-time and unstructured data, and fortifying governance and compliance controls. By doing so, they not only unlock greater AI performance and reliability but also create a proprietary data moat that protects their competitive position. In contrast, organisations that cling to ad-hoc, patchwork data environments will find their AI ambitions constrained - no matter how much they spend on algorithms.

References:

[1] aiireland.ie — https://aiireland.ie/2026/03/25/the-new-moat-why-proprietary-data-is-your-only-durable-competitive-advantage-in-ai/#:~:text=their%20cloud%20platforms,This%20is%20not%20a

[2] aiireland.ie — https://aiireland.ie/2026/03/25/the-new-moat-why-proprietary-data-is-your-only-durable-competitive-advantage-in-ai/#:~:text=models%20are%20becoming%20a%20commodity,The%20Model%20Race%20Is

[3] www.storagenewsletter.com — https://www.storagenewsletter.com/2026/07/02/couchbase-launches-the-ai-data-plane-the-operational-data-foundation-for-the-agentic-enterprise/#:~:text=StorageNewsletter,The%20AI%20Data

[4] inferenz.ai — https://inferenz.ai/blogs/databricks-data-ai-summit-2026-the-lakehouse-just-became-something-bigger/#:~:text=Data%20%2B%20AI%20Summit%202026%3A,produces%20a%20hundred%20summits%20that

[5] www.digitaltoday.co.kr — https://www.digitaltoday.co.kr/en/view/61144/microsoft-fabric-aims-to-be-data-platform-for-ai-agents-targeting-snowflake-and-databricks#:~:text=,and%20boundaries%20for%20AI%20agents

Key Statistics

- 72% of global IT leaders say a lack of real-time data infrastructure is stalling their efforts to scale AI ([www.confluent.io](https://www.confluent.io/press-release/2026-data-streaming-report/#:~:text=Calif.,examines%20the%20challenges%20that%20enterprises)).
- Only 6% of enterprise AI leaders consider their data infrastructure fully AI-ready ([www.cdata.com](https://www.cdata.com/company/press/state-of-ai-data-connectivity-report/#:~:text=Only%206,which%20draws%20on%20independently%20collected)).
- Gartner predicts 60% of AI projects lacking 'AI-ready' data will be abandoned by 2026 ([www.freevacy.com](https://www.freevacy.com/news/gartner/gartner-reveals-60-of-ai-projects-with-data-issues-will-fail/6155#:~:text=initiatives,Establish%20robust%20data%20governance)).
- Half of leading-edge companies report AI-driven ROI above 25%, compared to just 11% of early-stage firms ([venturebeat.com](https://venturebeat.com/orchestration/box-survey-why-enterprise-ai-leaders-are-outperforming-their-peers#:~:text=early,leading%20edge%20is%20the%20operating)).

KEY TAKEAWAY

The past week’s developments show that winning with AI means getting your data house in order. Organisations with connected, high-quality, well-governed data are achieving real ROI from AI, while those that don’t face stalled projects, compliance crackdowns, and competitive disadvantage.

Sources

AI ambitions at risk: 72% of IT leaders say poor infrastructure is stalling AI growth
https://www.confluent.io/press-release/2026-data-streaming-report/

Only 6% of AI Leaders Say Their Data Infrastructure Is Ready
https://www.cdata.com/company/press/state-of-ai-data-connectivity-report/

Box survey: Why enterprise AI leaders are outperforming their peers
https://venturebeat.com/orchestration/box-survey-why-enterprise-ai-leaders-are-outperforming-their-peers

Starbucks scraps AI inventory tool across North America
https://www.cnbc.com/2026/05/21/starbucks-scraps-ai-inventory-tool-across-north-america.html

Couchbase Launches the AI Data Plane, the Operational Data Foundation for the Agentic Enterprise
https://www.couchbase.com/press-releases/couchbase-launches-the-ai-data-plane-the-operational-data-foundation-for-the-agentic-enterprise/

The New Moat: Why Proprietary Data Is Your Only Durable Competitive Advantage in AI
https://aiireland.ie/2026/03/25/the-new-moat-why-proprietary-data-is-your-only-durable-competitive-advantage-in-ai/

EU AI Act Deadlines 2026-2027: Compliance Calendar + Fines
https://www.legiscope.com/blog/eu-ai-act-timeline-deadlines.html