LUMO INSIGHTS

Monday, 22 June 2026

Data strategy: AI’s make-or-break factor

Executive Summary

A spate of new research and industry moves this week underscores that the greatest gains - and risks - in AI now hinge on data. From infrastructure and quality to governance and ownership, leading enterprises are shifting focus to strengthening data foundations as the key to unlocking AI’s value.

The AI Data infrastructure divide

The past week has offered fresh evidence that robust data architecture is what separates AI leaders from laggards. Only 6% of enterprise AI leaders today say their data infrastructure is completely AI-ready, leaving a vast 'readiness gap' that has become one of the biggest constraints on AI progress (www.prnewswire.com [1]). In contrast, companies at the highest levels of AI maturity have all invested heavily in modern, centralised data infrastructure, while over half of those struggling with AI are hampered by fragmented, siloed data systems (www.prnewswire.com [2]). The result is a widening performance chasm: those with clean, well-integrated data can drive AI at scale, whereas those without it find themselves fighting data bottlenecks and fragmented deployments (hakkoda.io [3]) (hakkoda.io [4]).

Forward-looking organisations are accordingly shifting their spending from model development to fortifying data foundations. One recent industry report found that only 9% of enterprises now rank building new AI models as their top priority, while 83% are investing in unified data access and integration layers to fuel AI across the business (www.prnewswire.com [5]). Similarly, a Gartner survey revealed that companies successfully deploying AI allocate up to four times more of their annual revenue to data management, governance, and analytics than peers with poor AI results (www.gartner.com [6]). This rebalancing of investment towards data architecture signals a recognition that AI’s value is ultimately constrained by the quality, availability, and connectedness of the information it learns from.

As Amit Sharma, CEO of data connectivity firm CData, put it: the era of AI being limited by models is over - today, AI is constrained by data (www.prnewswire.com [7]). In practice, the organisations winning with AI 'aren’t the ones with the best algorithms; they’re the ones with connected, contextual, and semantically consistent data infrastructure' (www.prnewswire.com [8]). The takeaway for executives is clear: embracing modern, integrated data architecture - spanning real-time data sources, common standards, and end-to-end data governance - is becoming the decisive factor in AI competitiveness.

References:

[1] www.prnewswire.com — https://www.prnewswire.com/news-releases/cdata-study-finds-only-6-of-ai-leaders-believe-their-data-infrastructure-is-ready-for-ai-302629216.html#:~:text=article%20More%20than%20200%20AI,a%20central%20finding%20of%20CData

[2] www.prnewswire.com — https://www.prnewswire.com/news-releases/cdata-study-finds-only-6-of-ai-leaders-believe-their-data-infrastructure-is-ready-for-

ai-302629216.html#:~:text=https%3A%2F%2Fwww.cdata.com%2Flp%2Fai,is%20costing%20companies%20time%2C%20money
[3] hakkoda.io — https://hakkoda.io/state-of-data-2026/#:~:text=how%20organizations%20operate%2C%20compete%2C%20and,Key%20insights
[4] hakkoda.io — https://hakkoda.io/state-of-data-2026/#:~:text=infrastructure%20readiness%20is%20quietly%20determining,interoperability%20and%20scalability%20required%20to
[5] www.prnewswire.com — https://www.prnewswire.com/news-releases/cdata-study-finds-only-6-of-ai-leaders-believe-their-data-infrastructure-is-ready-for-ai-302629216.html#:~:text=Only%209,It%27s%20determined%20by%20the
[6] www.gartner.com — https://www.gartner.com/en/newsroom/press-releases/2026-04-16-gartner-says-organizations-with-successful-ai-initiatives-invest-up-to-four-times-more-in-data-and-analytics-foundations#:~:text=Context%20Foundations%20and%20Perceptive%20Intelligence,business%20and%20technology%20insights%20company
[7] www.prnewswire.com — https://www.prnewswire.com/news-releases/cdata-study-finds-only-6-of-ai-leaders-believe-their-data-infrastructure-is-ready-for-ai-302629216.html#:~:text=and%20competitive%20advantage.%20,with%20connected%2C%20contextual%2C%20and%20semantically
[8] www.prnewswire.com — https://www.prnewswire.com/news-releases/cdata-study-finds-only-6-of-ai-leaders-believe-their-data-infrastructure-is-ready-for-ai-302629216.html#:~:text=and%20competitive%20advantage.%20,with%20connected%2C%20contextual%2C%20and%20semantically

From bottleneck to breakthrough: quality & governance

Despite large investments in AI tools, many enterprises continue to hit the same wall: data problems. Gartner underscores this in a stark prediction that by 2026, 80% of organisations seeking to scale AI will fail to realise value because they have not modernised their data governance and integration architecture (optimumpartners.com [1]). Other research reinforces how common AI project failure has become. RAND Corporation estimates that more than 80% of AI projects never deliver their intended business value - roughly twice the failure rate of typical IT projects - and MIT found 95% of generative AI pilots show no measurable return on investment (www.pertamapartners.com [2]). The root causes behind these failures are strikingly consistent: unclear success metrics, disjointed legacy data silos, poor data quality, and absence of clear data ownership and governance to ensure trust in the results (www.pertamapartners.com [3]).

Real-world examples illustrate how flawed data can derail AI. In one case, MD Anderson Cancer Centre spent over $60 million on an AI-powered oncology system that ultimately had to be abandoned - the model had been trained on hypothetical patient data instead of real medical records (www.taleofdata.com [4]). Without accurate data, even a cutting-edge AI produced useless treatment recommendations. In another incident, Air Canada’s customer service chatbot provided a passenger with incorrect refund information, and a court held the airline liable - forcing it to honour the erroneous offer made by its AI agent (www.taleofdata.com [5]). Both cases show how poor data quality and oversight can turn AI from an asset into a liability.

Fortunately, forward-thinking organisations are turning data governance and quality from a barrier into a catalyst. By treating data as a strategic asset and enforcing rigorous standards for quality, metadata, and access, they ensure AI systems have reliable inputs from day one. Companies with mature AI governance are seeing tangible benefits - one industry survey linked strong governance practices to a 27% improvement in AI project efficiency and time-to-value (hakkoda.io [6]). In short, investing in data quality, lineage, and stewardship is not just about compliance; it’s about enabling faster innovation and greater confidence in AI-driven decisions.

References:

[1] optimumpartners.com — https://optimumpartners.com/insight/the-ai-data-bottleneck-why-data-pipelines-define-ai-success/#:~:text=access%2C%20trust%2C%20or%20understand,just%20use%20bad%20data%3B%20it
[2] www.pertamapartners.com — https://www.pertamapartners.com/insights/ai-project-failure-statistics-2026#:~:text=RAND%20Corporation%20estimates%20that%20more,measurable%20return%20to%20the%20income
[3] www.pertamapartners.com — https://www.pertamapartners.com/insights/ai-project-failure-statistics-2026#:~:text=estimates%2C%20more%20than%2080,not%20a%20precise%20percentage
[4] www.taleofdata.com — https://www.taleofdata.com/blog/why-ai-projects-fail-data-quality#:~:text=million%20investment%20from%20MD%20Anderson,severely%20diminished%20physician
[5] www.taleofdata.com — https://www.taleofdata.com/blog/why-ai-projects-fail-data-quality#:~:text=the%20project%20was%20finally%20abandoned,This%20case%20highlights%20the%20concrete
[6] hakkoda.io — https://hakkoda.io/state-of-data-2026/#:~:text=support%20autonomous%20systems,design%20architectures%20are

Data as a competitive Moat

As cutting-edge AI becomes broadly accessible, companies are grappling with a new reality: proprietary data - not algorithms - increasingly drives competitive advantage. Industry analysts note that today's powerful AI models are turning into 'strategic commodities' that offer little sustainable edge by themselves, since competitors can often access similar capabilities via cloud providers (aiireland.ie [1]). If every player can utilise the same state-of-the-art models, the differentiator becomes how effectively each organisation harnesses data that others cannot. In short, the one thing rivals can't easily copy is your trove of proprietary information - and in 2026, unique data has become the most durable competitive moat (aiireland.ie [2]).

Investors are already rewarding firms with strong data moats. According to Morningstar's analysis, companies most vulnerable to AI disruption (those lacking distinctive data assets) underperformed the most data-rich, 'AI-resilient' peers by nearly 26 percentage points in early 2026 (aiireland.ie [3]). This growing gap in market performance sends a clear signal: enterprises that have amassed large, high-quality datasets - from customer behaviour and product performance to supply chain intelligence or domain-specific research - can train more specialised, high-impact AI models and services that others struggle to match.

Unsurprisingly, leading organisations are doubling down on data as proprietary intellectual property. Many firms are bolstering data governance and security to control access to their most valuable datasets and prevent unwitting leaks that could erode their advantage. Some are even exploring new revenue streams by licensing or monetising data (where regulations permit) and acquiring data-rich companies to expand their knowledge base. For boards and CDOs, the mandate is clear: safeguarding and fully leveraging unique data assets is now essential to building and defending an AI-powered competitive edge.

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=strategic%20forecasting%2C%20are%20rapidly%20becoming,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=Moat%3A%20Why%20Proprietary%20Data%20Is,past%20three%20years%2C%20the%20boardroom

[3] aiireland.ie — https://aiireland.ie/2026/03/25/the-new-moat-why-proprietary-data-is-your-only-durable-competitive-advantage-in-ai/#:~:text=theoretical%20debate,for%20boards%20is%20stark%3A%20your

Building for real-time, integrated Data

To support AI's growing demands, enterprise technology providers are racing to remove data friction. One urgent focus is ensuring up-to-the-second information for AI-driven decisions. In a recent survey, 100% of AI leaders agreed that real-time data is essential for effective AI agents, yet 20% admitted their systems still lack real-time integration capabilities (www.prnewswire.com [1]). This 'data freshness' gap is forcing IT teams to overhaul legacy batch processes in favour of continuous data flows, so AI applications are never stuck using stale or siloed information.

The past week saw a major leap forward in data architecture designed for AI. At its Data + AI Summit, Databricks unveiled a new approach called Lakehouse Transactional/Analytical Processing (LTAP) that unifies transactional, analytical, streaming, and operational workloads on a single copy of data in a cloud data lake (www.storagenewsletter.com [2]). With one governed source of truth in the lake, enterprises can 'read, reason, and act' on real-time information without resorting to multiple databases or brittle extract-transform-load (ETL) pipelines (www.storagenewsletter.com [3]). In practical terms, AI models and analytics can now draw from the same fresh dataset simultaneously - a breakthrough that promises faster insights and more responsive, AI-powered operations.

Other major data platforms are following suit. Earlier this month, Snowflake announced an Interoperable Lakehouse architecture, highlighted by a new Horizon data catalogue that provides a consistent semantic and governance layer across the entire enterprise data estate (www.dbgurus.com.au [4]). The goal is to give both human users and AI systems a shared, trusted source of enterprise truth for queries and decisions. Traditional database vendors are also adapting: Oracle, for example, has rolled out an Autonomous AI Vector Database to embed vector search and retrieval into its cloud platform with full enterprise security and compliance support (blogs.oracle.com [5]). These trends all point in the same direction - building flexible, unified data architectures is now a strategic imperative for any company aiming to enable real-time, context-rich AI at scale.

References:

[1] www.prnewswire.com — https://www.prnewswire.com/news-releases/cdata-study-finds-only-6-of-ai-leaders-believe-their-data-infrastructure-is-ready-for-ai-302629216.html#:~:text=Real,Are%20Outpacing%20Traditional%20Software%20in
[2] www.storagenewsletter.com — https://www.storagenewsletter.com/2026/06/18/data-ai-summit-2026-databricks-launches-ltap-the-first-lake-transactional-analytical-processing-architecture/#:~:text=AI%20company%2C%20introduced%20Lake%20Transactional%2FAnalytical,to%20enlarge%20The%20New%20Data
[3] www.storagenewsletter.com — https://www.storagenewsletter.com/2026/06/18/data-ai-summit-2026-databricks-launches-ltap-the-first-lake-transactional-analytical-processing-architecture/#:~:text=and%20operational%20data%20on%20a,to%20enlarge%20The%20New%20Data
[4] www.dbgurus.com.au — https://www.dbgurus.com.au/the-agentic-first-data-lakehouse-how-ai-agents-are-rewriting-the-rules-of-enterprise-data-architecture/#:~:text=human%20interpretation,the%20core%20failure%20modes%20of
[5] blogs.oracle.com — https://blogs.oracle.com/database/announcing-oracle-autonomous-ai-vector-database-limited-availability#:~:text=help%20developers%20and%20their%20organizations,create%20new%20AI%20applications%20or

Regulation and the Data readiness mandate

The final driver placing data strategy at the top of the C-suite agenda is the changing regulatory environment. The European Union's expansive AI Act - the world's first comprehensive AI law - will begin enforcing key provisions in August 2026. Regulators have delayed some high-risk AI requirements until 2027, but this year's deadline still triggers new rules like mandatory transparency labels for AI-generated content and gives authorities broad power to audit and fine non-compliant organisations. In effect, any company operating in Europe (or handling EU consumer data) will soon need rigorous controls over its AI training data and models - from documented data lineage and bias testing to robust human oversight - or risk steep penalties.

This push for accountability is not limited to Europe. Over 80% of the world's population is now covered by some form of data protection law as countries from Brazil to China enact privacy and AI regulations (www.tjc-group.com [1]). Even in the United States, where federal AI rules remain in flux, regulators are applying existing consumer privacy and safety laws to AI deployments. In response, data sovereignty - keeping sensitive data and AI processes within specific jurisdictions - has become a top-of-mind concern for global enterprises. One recent survey found 77% of companies now factor a cloud provider's country-of-origin into AI vendor selection, and nearly three in five are choosing local or regional cloud options to better control their data and compliance risk (www.prnewswire.com [2]).

For CTOs and CDOs, these regulations make data readiness synonymous with business readiness. Compliance is increasingly intertwined with data architecture: organisations must ensure proper consent, quality, and governance of data used in AI systems to meet standards such as GDPR and the AI Act's data governance requirements (www.ailoitte.com [3]). Yet far from being just a legal headache, this rigorous approach to data can become a competitive advantage. Companies that get their data house in order - managing information ethically, transparently, and securely - won't just avoid fines; they'll also build greater trust with customers and regulators, paving the way for wider AI adoption and success.

References:

[1] www.tjc-group.com — https://www.tjc-group.com/blogs/data-privacy-in-2026-how-gdpr-compliance-landscape-is-evolving/

#:~:text=existing%20frameworks%2C%20the%20emergence%20of,term
[2] www.prnewswire.com — https://www.prnewswire.com/news-releases/from-ambition-to-activation-organizations-stand-at-the-untapped-edge-of-ais-potential-reveals-deloitte-survey-302666072.html#:~:text=organizations%20are%20taking%20note,and%20defense%20leading%20the%20way
[3] www.ailoitte.com — https://www.ailoitte.com/insights/eu-ai-act-gdpr-compliance-guide/#:~:text=GDPR%20Was%20the%20Beginning%3A%20How,AI%20Act%E2%80%99s%20data%20governance%20standards

Key Statistics

- Only 6% of enterprise AI leaders say their data infrastructure is fully ready for AI ([www.prnewswire.com](https://www.prnewswire.com/news-releases/cdata-study-finds-only-6-of-ai-leaders-believe-their-data-infrastructure-is-ready-for-ai-302629216.html#:~:text=article%20More%20than%20200%20AI,a%20central%20finding%20of%20CData)).
- More than 80% of AI projects fail to deliver their intended business value ([www.pertamapartners.com](https://www.pertamapartners.com/insights/ai-project-failure-statistics-2026#:~:text=RAND%20Corporation%20estimates%20that%20more,measurable%20return%20to%20the%20income)).
- Successful AI companies invest up to 4x more of revenue in data and analytics foundations than those with poor AI outcomes ([www.gartner.com](https://www.gartner.com/en/newsroom/press-releases/2026-04-16-gartner-says-organizations-with-successful-ai-initiatives-invest-up-to-four-times-more-in-data-and-analytics-foundations#:~:text=Context%20Foundations%20and%20Perceptive%20Intelligence,business%20and%20technology%20insights%20company)).
- Organisations with mature AI governance achieve about 27% faster and more efficient AI project outcomes ([hakkoda.io](https://hakkoda.io/state-of-data-2026/#:~:text=support%20autonomous%20systems,design%20architectures%20are)).
- 77% of companies factor a vendor's country-of-origin into AI decisions to ensure compliance with data regulations ([www.prnewswire.com](https://www.prnewswire.com/news-releases/from-ambition-to-activation-organizations-stand-at-the-untapped-edge-of-ais-potential-reveals-deloitte-survey-302666072.html#:~:text=organizations%20are%20taking%20note,and%20defense%20leading%20the%20way)).

KEY TAKEAWAY

AI leaders differentiate through superior data foundations. With new regulations looming and models becoming commodities, CTOs and CDOs must double down on data quality, governance and unified architectures to drive real AI value.

Sources

CData Study: Only 6% of AI Leaders' Data Infrastructure Is AI-Ready (PR Newswire, Dec 2025)
https://www.prnewswire.com/news-releases/cdata-study-finds-only-6-of-ai-leaders-believe-their-data-infrastructure-is-ready-for-ai-302629216.html

2026 State of Data Report – Hakkoda (Hybrid-By-Design Architecture & Governance)
https://hakkoda.io/state-of-data-2026/

AI Project Failure Statistics 2026 – Pertama Partners (June 2026)
https://www.pertamapartners.com/insights/ai-project-failure-statistics-2026

Why AI Projects Fail — The Data Quality Crisis (Tale of Data, 2025)
https://www.taleofdata.com/blog/why-ai-projects-fail-data-quality

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

Data + AI Summit 2026: Databricks Launches LTAP – The First Lake Transactional/Analytical Processing Architecture (StorageNewsletter, June 2026)
https://www.storagenewsletter.com/2026/06/18/data-ai-summit-2026-databricks-launches-ltap-the-first-lake-transactional-analytical-processing-architecture/

The Agentic-First Data Lakehouse: How AI Agents Are Rewriting the Rules of Enterprise Data Architecture (Snowflake

Summit 2026 – DB Gurus)
https://www.dbgurus.com.au/the-agentic-first-data-lakehouse-how-ai-agents-are-rewriting-the-rules-of-enterprise-data-architecture/

Announcing Oracle Autonomous AI Vector Database – Oracle Database Blog (Mar 2026)
https://blogs.oracle.com/database/post/announcing-oracle-autonomous-ai-vector-database-limited-availability

Data Privacy in 2026: How the GDPR Compliance Landscape Is Evolving (TJC Group, 2026)
https://www.tjc-group.com/blogs/data-privacy-in-2026-how-gdpr-compliance-landscape-is-evolving/

From Ambition to Activation – State of AI in the Enterprise 2026 (Deloitte AI Institute, Jan 2026)
https://www.prnewswire.com/news-releases/from-ambition-to-activation-organizations-stand-at-the-untapped-edge-of-ais-potential-reveals-deloitte-survey-302666072.html

GDPR Was the Beginning: How the EU AI Act Changes Everything (Ailoitte, 2026)
https://www.ailoitte.com/insights/eu-ai-act-gdpr-compliance-guide/

Gartner Press Release (Apr 2026): Successful AI Initiatives Invest 4x More in Data and Analytics Foundations
https://www.gartner.com/en/newsroom/press-releases/2026-04-16-gartner-says-organizations-with-successful-ai-initiatives-invest-up-to-four-times-more-in-data-and-analytics-foundations