LUMO INSIGHTS
Monday, 18 May 2026

Why data strategy is the deciding factor in the AI race

Executive Summary

Over the past week, multiple new studies and industry moves have underscored that data readiness - not model prowess - is the critical determinant of AI success (www.fivetran.com [1]). Organisations with high-quality, well-governed, and connected data foundations are achieving far greater AI-driven value, while those lacking in data quality and integration are seeing projects stall or fail at alarming rates (workos.com [2]). For the C-suite, the message is clear: only by prioritising data strategy (from quality and governance to architecture and compliance) can businesses turn AI investments into sustainable ROI (towardsdatascience.com [3]).

References:

[1] www.fivetran.com — https://www.fivetran.com/press/fivetran-launches-2026-agentic-ai-readiness-index-revealing-gap-between-enterprise-investment-and-data-preparedness-for-agentic-ai#:~:text=findings%20show%20that%20only%2015,Pacific%2C%20the%20index
[2] workos.com — https://workos.com/blog/why-most-enterprise-ai-projects-fail-patterns-that-work#:~:text=Perplexity%20The%20statistics%20are%20sobering,and%20security%20risks%20as%20the
[3] towardsdatascience.com — https://towardsdatascience.com/the-2026-data-mandate-is-your-governance-architecture-a-fortress-or-a-liability/#:~:text=In%202026%2C%20European%20companies%20can,tax%20to%20a%20legal%20liability

The data readiness divide

A stark new global index has exposed how far behind most companies are in data readiness for AI. The 2026 “agentic AI readiness” survey finds that only 15% of organisations have the necessary AI-ready data foundation to reliably support advanced AI, even though nearly 60% have already invested millions in AI technology (www.fivetran.com [1]). In fact, 41% of enterprises are already deploying autonomous AI agents in production despite significant gaps in data reliability, governance, and interoperability (www.fivetran.com [2]).

For many, this gap between ambition and preparedness has tangible consequences. An analysis by S&P Global found 42% of companies abandoned most of their AI projects in 2025 - a dramatic jump from just 17% the year prior (workos.com [3]). Similarly, a recent MIT study estimates up to 95% of AI projects ultimately fail to deliver on their promises, with poor data quality and integration the chief culprits (www.forbes.com [4]) (www.forbes.com [5]). Companies burden models with incomplete, biased, or outdated information, leading to brittle models that struggle under real-world conditions (www.forbes.com [6]).

Meanwhile, a small minority of AI leaders are pulling ahead by treating data as a strategic asset rather than an afterthought. According to PwC’s new AI Performance Study, nearly 74% of all AI economic value is being captured by just 20% of companies (www.pwc.com [7]). These top performers aren’t simply experimenting with more algorithms - they use AI to drive growth and “back that ambition with the foundations that make AI scalable and reliable,” including strong data governance and trust infrastructures (www.pwc.com [8]). In other words, as IBM’s CEO Arvind

Krishna put it, the enterprises winning with AI 'are not deploying more AI - they're redesigning how their business operates' (newsroom.ibm.com [9]), starting with robust data foundations.

References:

[1] www.fivetran.com — https://www.fivetran.com/press/fivetran-launches-2026-agentic-ai-readiness-index-revealing-gap-between-enterprise-investment-and-data-preparedness-for-agentic-ai#:~:text=findings%20show%20that%20only%2015,Pacific%2C%20the%20index

[2] www.fivetran.com — https://www.fivetran.com/press/fivetran-launches-2026-agentic-ai-readiness-index-revealing-gap-between-enterprise-investment-and-data-preparedness-for-agentic-ai#:~:text=that%20were%20never%20designed%20for,of%20data%20leaders%20say

[3] workos.com — https://workos.com/blog/why-most-enterprise-ai-projects-fail-patterns-that-work#:~:text=Perplexity%20The%20statistics%20are%20sobering,and%20security%20risks%20as%20the

[4] www.forbes.com — https://www.forbes.com/sites/garydrenik/2025/10/15/why-95-of-ai-projects-fail-and-how-better-data-can-change-that/#:~:text=predictive%20systems%20are%20keeping%20vital,So%20Many%20AI%20Projects%20Fail

[5] www.forbes.com — https://www.forbes.com/sites/garydrenik/2025/10/15/why-95-of-ai-projects-fail-and-how-better-data-can-change-that/#:~:text=incomplete%2C%20biased%2C%20or%20stale%20datasets%2C,similarly%20cautioned%20that%20data%20provenance

[6] www.forbes.com — https://www.forbes.com/sites/garydrenik/2025/10/15/why-95-of-ai-projects-fail-and-how-better-data-can-change-that/#:~:text=%E2%80%94meaning%20it%20is%20accurate%2C%20representative%2C,to%20AI%20trustworthiness%20as%20the

[7] www.pwc.com — https://www.pwc.com/gx/en/news-room/press-releases/2026/pwc-2026-ai-performance-study.html#:~:text=they%20are%20seeing%20from%20AI,leaders%20and%20the%20majority%20of

[8] www.pwc.com — https://www.pwc.com/gx/en/news-room/press-releases/2026/pwc-2026-ai-performance-study.html#:~:text=businesses%20still%20stuck%20in%20pilot,Global%20Chief%20AI%20Officer%2C%20PwC

[9] newsroom.ibm.com — https://newsroom.ibm.com/2026-05-05-think-2026-ibm-delivers-the-blueprint-for-the-ai-operating-model-as-the-ai-divide-widens#:~:text=invested%20heavily%20in%20AI%2C%20but,Running%20AI%20in%20the

Modern architecture: from silos to real-time

Many companies are discovering that legacy data architectures are ill-suited for AI's demands. Data often remains locked in silos across different applications and clouds, forcing AI systems to work with incomplete or stale information. Traditional approaches bolt on multiple specialised data stores - from data warehouses and lakes to real-time event streams and new vector databases for AI search - each handling a fragment of enterprise information. This patchwork creates complexity and delay: teams must build sync pipelines to move and reconcile data between systems, which introduces latency, governance gaps, and higher costs (venturebeat.com [1]). The result is that AI agents frequently break at the data layer, hitting stale context, inconsistent access controls, and "fragmentation fatigue" as soon as they attempt to scale (venturebeat.com [2]).

To address these issues, leading providers are introducing platforms that make enterprise data both unified and instantly accessible to AI. Google, for example, has developed a next-generation cross-cloud data "lakehouse" that replaces batch processing with continuous data streams and a semantic context layer (cloud.google.com [3]). By keeping data in sync across different cloud environments without manual ETL, this open-standards approach can give AI real-time situational awareness. Google estimates an impressive 117% return on investment within six months for organisations adopting its cross-cloud lakehouse model (cloud.google.com [4]).

Established enterprise software players are also doubling down on unified data foundations. At its Knowledge 2026 conference this month, ServiceNow introduced a "real-time data foundation" designed to deliver live, trusted information directly into AI-driven workflows (newsroom.servicenow.com [5]). This new "Context Engine" pulls signals from across IT and business systems (from workflow logs to operational metrics) and integrates them under a common governance layer, ensuring every AI action is grounded in up-to-the-minute context (newsroom.servicenow.com [6]). By embedding data quality controls and automated lineage tracking, ServiceNow's approach tackles the root cause of many AI project failures: data fragmented across disconnected sources and left ungoverned at the moment of decision, which leads to shallow recommendations instead of reliable execution (newsroom.servicenow.com [7]).

Database giant Oracle is taking a different route by bringing AI's data needs under one roof. In late March, Oracle announced new "agentic AI" features for its flagship database, including a Unified Memory engine that can process vectors (for language model embeddings), JSON documents, graph data and traditional relational records together in one ACID-compliant system (venturebeat.com [8]). Alongside this, Oracle launched a fully managed Vector Database service built on the same engine, giving developers an easy on-ramp that can seamlessly upgrade to the complete converged database as their AI workloads grow (venturebeat.com [9]). Oracle's bet is that converging these technologies will reduce latency and eliminate painful data copying - automatically keeping AI models in sync with the latest enterprise data and enforcing security policies consistently. Industry analysts note that managing separate vector, graph and relational stores for one AI application is a 'DevOps nightmare' for data teams, and Oracle's all-in-one architecture is explicitly aimed at ending that fragmentation burden (venturebeat.com [10]).

References:

[1] venturebeat.com — https://venturebeat.com/data/oracle-converges-the-ai-data-stack-to-give-enterprise-agents-a-single#:~:text=Image%20generated%20by%20VentureBeat%20with,built%20around%20a%20direct%20architectural

[2] venturebeat.com — https://venturebeat.com/data/oracle-converges-the-ai-data-stack-to-give-enterprise-agents-a-single#:~:text=Enterprise%20agent%20deployments%20are%20not,Dickens

[3] cloud.google.com — https://cloud.google.com/blog/products/data-analytics/the-future-of-data-lakehouse-for-the-agentic-era#:~:text=engineered%20for%20the%20era%20of,performance%2C%20scale%2C%20governance%2C%20and%20multimodal

[4] cloud.google.com — https://cloud.google.com/blog/products/data-analytics/the-future-of-data-lakehouse-for-the-agentic-era#:~:text=approach%20can%20deliver%20an%20estimated,the%20silos%20between%20our%20data

[5] newsroom.servicenow.com — https://newsroom.servicenow.com/press-releases/details/2026/ServiceNow-launches-the-real-time-data-foundation-that-puts-autonomous-AI-to-work-across-the-enterprise/default.aspx#:~:text=enterprise%20intelligence,and%20applying%20a

[6] newsroom.servicenow.com — https://newsroom.servicenow.com/press-releases/details/2026/ServiceNow-launches-the-real-time-data-foundation-that-puts-autonomous-AI-to-work-across-the-enterprise/default.aspx#:~:text=fragmented%20across%20disconnected%20systems%20and,more%20accurate%20the%20more%20it

[7] newsroom.servicenow.com — https://newsroom.servicenow.com/press-releases/details/2026/ServiceNow-launches-the-real-time-data-foundation-that-puts-autonomous-AI-to-work-across-the-enterprise/default.aspx#:~:text=siloed%20across%20systems%20that%20has,workflow%20data%2C%20analytics%20insights%2C%20and

[8] venturebeat.com — https://venturebeat.com/data/oracle-converges-the-ai-data-stack-to-give-enterprise-agents-a-single#:~:text=that%27s%20not%20true.,By%20having%20the

[9] venturebeat.com — https://venturebeat.com/data/oracle-converges-the-ai-data-stack-to-give-enterprise-agents-a-single#:~:text=in%20a%20single%20query,same%20standard%20API%20call%20you

[10] venturebeat.com — https://venturebeat.com/data/oracle-converges-the-ai-data-stack-to-give-enterprise-agents-a-single#:~:text=agents%20are%20effectively%20stateless%20while,one%20agent%20is%20a%20DevOps

Proprietary data as the new moat

As AI models become ubiquitous and commoditised, companies are increasingly looking to their data as the source of competitive advantage. Proprietary, high-quality datasets - especially those rich in context and unique to a business - are becoming strategic Intellectual Property. One industry analyst recently likened these coveted datasets to the 'rare earth elements' of the AI era: scarce, hard-to-mine resources that have outsized importance in enabling powerful AI outcomes (www.forbes.com [1]). A leading bank's data chief even observed that we've 'already run out of data' in the public domain, underscoring that future AI breakthroughs will rely on proprietary, hard-to-replicate data sources (www.forbes.com [2]).

Real-world market developments are bearing out this "data moat" strategy. Investors are pouring capital into established firms and startups that have access to unique data streams, whether from customers, sensors, or industry-specific sources. Conversely, AI companies without their own data are hitting walls. In the first quarter of 2026, 93% of AI-focused startups that failed fell victim to lacking AI-ready data and governance - despite collectively raising $15 /billion in funding (www.czconsultants.com [3]). These included many so-called "AI wrapper" startups, which simply layered generic models over public data and ultimately collapsed due to thin margins and no defensible data asset to differentiate them (www.czconsultants.com [4]). Analysts predict that in 2026

only those AI players with strong, domain-specific data moats will thrive, while others will struggle to prove ROI and retain market share (www.czconsultants.com [5]).

For incumbents, this means doubling down on data as a strategic asset. Leaders are exploring ways to expand and protect their proprietary data - whether through partnerships to pool industry data, improved data collection from customers, or acquiring data-rich businesses. The objective is to create self-reinforcing advantages: the more unique data a company gathers and learns from, the smarter its AI systems become, widening the gap against competitors. In effect, data is turning into an IP asset and protective moat around AI initiatives, one that can safeguard market position in ways that off-the-shelf algorithms cannot (www.forbes.com [6]).

References:
[1] www.forbes.com — https://www.forbes.com/sites/garydrenik/2026/02/10/ais-real-bottleneck-isnt-algorithms-its-the-rare-earths-of-data/#:~:text=performance%20collapses,see%20AI%20systems%20that%20look
[2] www.forbes.com — https://www.forbes.com/sites/garydrenik/2026/02/10/ais-real-bottleneck-isnt-algorithms-its-the-rare-earths-of-data/#:~:text=2025%2C%20Goldman%20Sachs%E2%80%99%20chief%20data,to
[3] www.czconsultants.com — https://www.czconsultants.com/page/ai-startups-93-failure-2026-governance-gap#:~:text=funding%20wipeout%20March%2024%2C%202026,90%25%20of%20AI%20projects%20risk
[4] www.czconsultants.com — https://www.czconsultants.com/page/ai-startups-93-failure-2026-governance-gap#:~:text=primarily%20driven%20by%20inadequate%20data,correction%20phase%2C%20with%20predictions%20of
[5] www.czconsultants.com — https://www.czconsultants.com/page/ai-startups-93-failure-2026-governance-gap#:~:text=unproven%20unit%20economics%2C%20as%20seen,due%20to%20absent%20governance%2C%20monitoring
[6] www.forbes.com — https://www.forbes.com/councils/forbescommunicationscouncil/2025/02/06/proprietary-data-is-the-new-economic-moat-not-ai/#:~:text=retail%20media%20offerings%2C%20a%20%24100,into%20emerging%20models%20for%20aggregating

Governance and ethics move to centre stage

With great data power comes great responsibility - and oversight. As AI projects scale, any weakness in data governance or quality becomes a multiplying liability rather than a minor IT issue. One expert warns that if you feed AI 'noise,' it will simply 'scale confusion,' whereas feeding it 'clarity' will 'scale intelligence' (www.prnewswire.com [1]). This reality is elevating data governance from an IT task to a boardroom priority. In a recent global survey, 40.9% of IT leaders said improving data governance is a top priority for 2026, even above specific AI initiatives (www.prnewswire.com [2]). Their goal: ensure the data fuelling AI is accurate, secure, and trustworthy - so that AI-driven decisions are as sound as the data behind them.

Regulators are also raising the bar for data management in AI. The European Union’s sweeping AI Act, set to take effect in 2026, will require companies deploying “high-risk AI” to prove the quality, transparency, and provenance of their training data (towardsdatascience.com [3]). Combined with existing privacy laws like GDPR and new data localisation rules around the world, 'messy data' is no longer just a technical debt - it’s now a legal liability (towardsdatascience.com [4]). Organisations must document where data comes from, how it’s used in AI models, and ensure it respects individual rights and sovereignty requirements. Failing to do so can result in fines, reputational damage, and the inability to deploy AI systems in key markets. On the flip side, companies that get data governance right are seeing direct benefits: for instance, those leading in AI are 1.7 times more likely to have a responsible-AI framework and dedicated governance boards, and their employees are twice as likely to trust AI outputs (www.pwc.com [5]). In the coming months, we can expect the smartest CDOs and CTOs to tighten data ownership, quality control, and compliance checks - not to slow down innovation, but to enable it at scale.

References:
[1] www.prnewswire.com — https://www.prnewswire.com/news-releases/data-priorities-2026-ai-adoption-exposes-gaps-in-data-quality-governance-and-literacy-says-info-tech-research-group-in-new-report-302672864.html#:~:text=across%20data%20programs.%20,Priorities%20for%202026%20To%20help
[2] www.prnewswire.com — https://www.prnewswire.com/news-releases/data-priorities-2026-ai-adoption-exposes-gaps-in-data-quality-governance-and-literacy-says-info-tech-research-group-in-new-report-302672864.html#:~:text=emphasizes%20that%20AI%20

and%20automation,report%20also%20reflects%20broader%20pressures

[3] towardsdatascience.com — https://towardsdatascience.com/the-2026-data-mandate-is-your-governance-architecture-a-fortress-or-a-liability/#:~:text=,The%20Quality%20%26%20Ethics%20Mandate

[4] towardsdatascience.com — https://towardsdatascience.com/the-2026-data-mandate-is-your-governance-architecture-a-fortress-or-a-liability/#:~:text=In%202026%2C%20European%20companies%20can,tax%20to%20a%20legal%20liability

[5] www.pwc.com — https://www.pwc.com/gx/en/news-room/press-releases/2026/pwc-2026-ai-performance-study.html#:~:text=%281,approach%2C%20the%20performance%20gap%20between

Key Statistics

- Only 15% of organisations are fully prepared for AI, despite nearly 60% investing millions ([www.fivetran.com](https://www.fivetran.com/press/fivetran-launches-2026-agentic-ai-readiness-index-revealing-gap-between-enterprise-investment-and-data-preparedness-for-agentic-ai#:~:text=findings%20show%20that%20only%2015,Pacific%2C%20the%20index)).
- Nearly 3/4 of AI's economic value is being captured by the top 20% of companies ([www.pwc.com](https://www.pwc.com/gx/en/news-room/press-releases/2026/pwc-2026-ai-performance-study.html#:~:text=they%20are%20seeing%20from%20AI,leaders%20and%20the%20majority%20of)).
- 42% of companies abandoned most of their AI initiatives in 2025, up from 17% in 2024 ([workos.com](https://workos.com/blog/why-most-enterprise-ai-projects-fail-patterns-that-work#:~:text=Perplexity%20The%20statistics%20are%20sobering,and%20security%20risks%20as%20the)).
- 40.9% of IT leaders cite improving data governance as a top priority for 2026 ([www.prnewswire.com](https://www.prnewswire.com/news-releases/data-priorities-2026-ai-adoption-exposes-gaps-in-data-quality-governance-and-literacy-says-info-tech-research-group-in-new-report-302672864.html#:~:text=emphasizes%20that%20AI%20and%20automation,report%20also%20reflects%20broader%20pressures)).

KEY TAKEAWAY

AI initiatives live or die by data. Leaders must invest in clean, connected, well-governed data and modern architectures now. Without a solid data foundation, AI projects will stall - and regulators are watching.

Sources

Fivetran Launches 2026 Agentic AI Readiness Index, Revealing Gap Between Enterprise Investment and Data Preparedness for Agentic AI
https://www.fivetran.com/press/fivetran-launches-2026-agentic-ai-readiness-index-revealing-gap-between-enterprise-investment-and-data-preparedness-for-agentic-ai

Three-quarters of AI's economic gains are being captured by just 20% of companies – with the leading companies focused on growth, not just productivity (PwC 2026 AI Performance Study)
https://www.pwc.com/gx/en/news-room/press-releases/2026/pwc-2026-ai-performance-study.html

ServiceNow launches the real-time data foundation that puts autonomous AI to work across the enterprise (ServiceNow Knowledge 2026 Press Release)
https://newsroom.servicenow.com/press-releases/details/2026/ServiceNow-launches-the-real-time-data-foundation-that-puts-autonomous-AI-to-work-across-the-enterprise/default.aspx

The future of data lakehouse: Open and interoperable for the agentic era (Google Cloud Blog)
https://cloud.google.com/blog/products/data-analytics/the-future-of-data-lakehouse-for-the-agentic-era

Oracle converges the AI data stack to give enterprise agents a single version of truth (VentureBeat)
https://venturebeat.com/data/oracle-converges-the-ai-data-stack-to-give-enterprise-agents-a-single-version-of-truth

AI's Real Bottleneck Isn't Algorithms, It's The Rare Earths Of Data (Forbes)
https://www.forbes.com/sites/garydrenik/2026/02/10/ais-real-bottleneck-isnt-algorithms-its-the-rare-earths-of-data/

Data Priorities 2026: AI Adoption Exposes Gaps in Data Quality, Governance, and Literacy (Info-Tech Research Group report)
https://www.prnewswire.com/news-releases/data-priorities-2026-ai-adoption-exposes-gaps-in-data-quality-governance-and-literacy-says-info-tech-research-group-in-new-report-301732725.html

New data reveals AI governance gap between policy and practice, creating ESG risks (Thomson Reuters Foundation)

https://www.thomsonreuters.com/en-us/posts/sustainability/ai-governance-gap-esg-risks/

Why most enterprise AI projects fail — and the patterns that actually work (WorkOS blog)
https://workos.com/blog/why-most-enterprise-ai-projects-fail-patterns-that-work

Proprietary Data Is The New Economic Moat, Not AI (Forbes)
https://www.forbes.com/sites/forbescommunicationscouncil/2025/02/06/proprietary-data-is-the-new-economic-moat-not-ai/