LUMO INSIGHTS
Monday, 21 September 2026

Data foundations: the deciding factor in enterprise AI success

Executive Summary

A wave of new reports and deals this week highlights a hard truth: the greatest hurdle in scaling artificial intelligence is not the algorithms or models - it's the state of enterprise data. From infrastructure and quality issues stalling projects to multi-billion-dollar bets on proprietary data, organisations are being forced to confront data strategy as the linchpin of AI success. Meanwhile, regulators are stepping in, making robust data governance a must-have for any AI initiative. Here's what happened and what it means for the C-suite.

Data infrastructure: the new AI bottleneck

Nearly half of AI initiatives are running into roadblocks because of data shortcomings, according to new industry findings (www.confluent.io [1]). In a global survey of 4,625 IT leaders, 72% reported encountering at least three major challenges when trying to scale AI - with insufficient real-time data infrastructure, data integration woes, and poor data quality among the top culprits (www.confluent.io [2]) (www.confluent.io [3]). In other words, the limiting factor for advanced AI is often not the model itself, but the pipelines and platforms feeding it reliable, accessible information.

These data bottlenecks are forcing a shift in how companies approach AI. Many organisations have been surprised to learn that even the most powerful algorithms cannot compensate for fragmented or low-quality data. Analysts note that at least half of generative AI projects are abandoned after the proof-of-concept stage due to issues like disorganized data, unclear ownership, and insufficient risk controls (www.ibm.com [4]). The message is clear: without a modern, well-integrated data architecture, promising AI prototypes simply can't transition into scalable, real-world solutions.

To overcome this, forward-thinking enterprises are re-architecting their data foundations for AI. They are embracing modern "lakehouse" designs and real-time streaming pipelines that unify data across silos, ensuring AI has timely, holistic information. Vendors are also responding with new capabilities aimed at eliminating data friction. For example, Databricks just rolled out a feature allowing companies to perform both traditional SQL queries and vector similarity searches in one platform, so AI applications can retrieve semantic context and structured data seamlessly without separate search systems (docs.databricks.com [5]). By investing in such integrated data infrastructure, businesses hope to remove the final technical hurdles and let AI solutions operate at full speed.

References:

[1] www.confluent.io — https://www.confluent.io/press-release/2026-data-streaming-report/#:~:text=investment%20in%20AI%20itself,production%2C%20with%20the%20majority%20experiencing
[2] www.confluent.io — https://www.confluent.io/press-release/2026-data-streaming-report/#:~:text=investment%20in%20AI%20itself,production%2C%20with%20the%20majority%20experiencing
[3] www.confluent.io — https://www.confluent.io/press-release/2026-data-streaming-report/#:~:text=encountered%20at%20least%20three%20challenges,production%2C%20with%20the%20majority%20experiencing
[4] www.ibm.com — https://www.ibm.com/think/insights/why-most-enterprise-ai-projects-stall-before-

scale#:~:text=don%E2%80%99t%20fail%20because%20of%20their,progress%20can%20slow%20and%20often
[5] docs.databricks.com — https://docs.databricks.com/aws/en/release-notes/lakebase/#:~:text=September%2018%2C%202026,without%20a%20separate%20search%20stack

The AI value gap: Data quality at fault

The vast majority of companies have yet to see tangible value from their AI efforts, and poor data quality is a primary reason why. A research collaboration by Accenture and Carnegie Mellon University's Software Engineering Institute revealed a startling statistic: 95% of organisations report no measurable return on their AI investments to date (www.sei.cmu.edu [1]). Only an elite 8% of firms have successfully scaled AI across the enterprise and realised significant benefits at scale (www.sei.cmu.edu [2]). Crucially, the study finds technology isn't the root problem - it's deficiencies in data readiness and execution that are holding businesses back (www.sei.cmu.edu [3]).

This is corroborated by other surveys showing that most AI projects struggle with foundational data issues. In one Harvard Business Review - Cloudera study, 73% of executives said their organisations need to prioritise AI data quality more, and a similar proportion admitted that just collecting, processing, and preparing data for AI is a major challenge (www.cloudera.com [4]). The top obstacles cited were telling: 56% pointed to siloed data and integration difficulties, 44% to a lack of clear data strategy, and 41% to data quality or bias problems - notably outranking concerns like regulatory constraints (www.cloudera.com [5]). In short, many enterprises are learning that AI projects often fail not because the algorithms are flawed, but because the input data is.

Concrete examples from the field illustrate this pattern. Consider a global bank that piloted an AI agent to automate regulatory reporting: the system could rapidly generate compliance reports from financial data and initially impressed stakeholders with its potential. However, the pilot never reached full deployment, as it depended on manually curated datasets and required human validation before outputs could flow into real workflows (www.ibm.com [6]). The underlying data architecture couldn't support the automation at scale, causing the project to stall despite a successful model. Such outcomes are prompting leadership teams to confront "data debt" - the accumulated shortfalls in data quality, integration, and governance - as a critical barrier to realising AI's promised ROI.

References:

[1] www.sei.cmu.edu — https://www.sei.cmu.edu/news/sei-and-accenture-release-ai-adoption-maturity-model-to-help-organizations-scale-ai-with-predictable-outcomes/#:~:text=and%20build%20a%20clear%20roadmap,Mismatched%20expectations%2C%20misaligned%20applications%2C%20and
[2] www.sei.cmu.edu — https://www.sei.cmu.edu/news/sei-and-accenture-release-ai-adoption-maturity-model-to-help-organizations-scale-ai-with-predictable-outcomes/#:~:text=are%20rising%2C%20but%2095%20percent,Mismatched%20expectations%2C%20misaligned%20applications%2C%20and
[3] www.sei.cmu.edu — https://www.sei.cmu.edu/news/sei-and-accenture-release-ai-adoption-maturity-model-to-help-organizations-scale-ai-with-predictable-outcomes/#:~:text=and%20build%20a%20clear%20roadmap,Mismatched%20expectations%2C%20misaligned%20applications%2C%20and
[4] www.cloudera.com — https://www.cloudera.com/about/news-and-blogs/press-releases/2026-03-05-only-7-percent-of-enterprises-say-their-data-is-completely-ready-for-ai-according-to-new-report-from-cloudera-and-harvard-business-review-analytic-services-reveals.html#:~:text=Perception%20Issues%20Despite%20widespread%20AI,future%20initiative%3B%20it%20is%20a
[5] www.cloudera.com — https://www.cloudera.com/about/news-and-blogs/press-releases/2026-03-05-only-7-percent-of-enterprises-say-their-data-is-completely-ready-for-ai-according-to-new-report-from-cloudera-and-harvard-business-review-analytic-services-reveals.html#:~:text=percentage%20report%20that%20their%20organization,future%20initiative%3B%20it%20is%20a
[6] www.ibm.com — https://www.ibm.com/think/insights/why-most-enterprise-ai-projects-stall-before-scale#:~:text=reporting,of

Leaders and laggards: Data as competitive moat

As the AI landscape matures, one factor increasingly separates winners from losers: how well companies leverage and protect their data. In an environment where sophisticated AI models and tools are becoming widely available (and the cost of advanced models is plummeting), proprietary data is now the key competitive differentiator (authority-journal.com [1]). In other words, when algorithms are no longer scarce, the unique data that organisations possess - customer interactions, operational insights, domain-specific datasets - becomes their AI advantage.

This was made evident by a headline-grabbing deal in the past week. Salesforce, a leader in enterprise software, was reported to be negotiating a $2 /billion acquisition of an AI startup called Listen Labs - a remarkable 67-times revenue multiple that stunned industry observers (enterprisedna.co [2]). The rationale was clear: Listen Labs isn't just another AI tool, it's a rich repository of customer research data. The three-year-old firm has built a 50-million-person panel and amassed over a million AI-mediated customer interviews, a trove of insights that Salesforce's own systems cannot duplicate (www.techtimes.com [3]). Paying a hefty premium for this startup underscores how urgently established companies want to secure unique data assets as "fuel" for their AI engines.

It's not just Salesforce. The most advanced AI adopters in every sector have been those that treat data as strategic intellectual property. Gartner's latest analysis indicates top AI performers invest up to four times more (as a percentage of revenue) in data quality, governance, and talent than their less advanced peers. That data-centric investment is paying off: organisations with mature, AI-ready data foundations have achieved as much as 65% greater improvements in revenue growth and cost reductions from AI initiatives. Companies that lag on AI often find themselves lacking the clean, integrated, and comprehensive data needed to train models effectively - and they risk falling further behind as leaders' data advantages compound over time.

References:

[1] authority-journal.com — https://authority-journal.com/the-proprietary-data-moat-who-actually-wins-when-ai-gets-cheap/#:~:text=As%20foundation%20model%20costs%20collapse,most%20are%20structuring%20it%20wrong

[2] enterprisedna.co — https://enterprisedna.co/resources/news/salesforce-listen-labs-2b-ai-customer-research-september-2026/#:~:text=Salesforce%20is%20in%20talks%20to,powered%20customer%20intelligence

[3] www.techtimes.com — https://www.techtimes.com/articles/327263/20260910/salesforce-eyes-2b-listen-labs-buy-fin-deal-closes-chasing-ai-data-moat.htm#:~:text=an%20AI%20market%20that%20has,who%20met%20at%20Harvard%20%E2%80%94

Governance and compliance: Data in the spotlight

Around the world, regulators are raising the stakes for data management in AI. In the past, enterprises could get by with aspirational principles and internal AI ethics pledges. Now, new laws are making robust data governance and documentation a requirement, not a choice (www.dataversity.net [1]). As of August, the European Union's AI Act has moved from theory to enforcement, imposing binding obligations on providers of high-risk and general-purpose AI systems (www.dataversity.net [2]). Businesses deploying advanced AI in the EU must be prepared to demonstrate exactly how their models are trained, how data quality and bias are managed, and how decisions and impacts are tracked and audited (www.dataversity.net [3]). In effect, AI oversight is shifting from the lab to the boardroom and compliance office.

Meanwhile, in the United States, where federal AI regulations remain in flux, state-level initiatives are picking up the slack. For example, California's governor faces a September 30 deadline to sign or veto a new "Frontier AI" law (SB 1047) aimed at the most powerful AI systems. If enacted, this law will require companies deploying very large AI models (those exceeding a high compute and investment threshold) to implement strict measures such as audited development logs and safety impact assessments. These moves come on top of existing data privacy regimes like GDPR, putting additional pressure on companies to get their data houses in order. The takeaway for organisations is clear: being an AI leader now means excelling not only at innovation, but also at data stewardship and regulatory compliance. In 2026, Chief Data Officers and CTOs are prioritising enterprise data architecture reviews, iron clad data governance frameworks, and transparent data usage policies to ensure their AI ambitions can move forward unimpeded.

References:
[1] www.dataversity.net — https://www.dataversity.net/articles/ai-governance-in-2026-is-your-organization-ready/#:~:text=Europe%2C%20the%20EU%20AI%20Act,how%20incidents%20were%20handled%2C%20and
[2] www.dataversity.net — https://www.dataversity.net/articles/ai-governance-in-2026-is-your-organization-ready/#:~:text=Europe%2C%20the%20EU%20AI%20Act,wasn%E2%80%99t%20philosophical%20alignment%2C%20but%20enforceability
[3] www.dataversity.net — https://www.dataversity.net/articles/ai-governance-in-2026-is-your-organization-ready/#:~:text=Europe%2C%20the%20EU%20AI%20Act,how%20incidents%20were%20handled%2C%20and

Key Statistics

- 72% - Share of IT leaders who report facing at least three major data infrastructure or quality challenges when trying to scale AI initiatives ([www.confluent.io](https://www.confluent.io/press-release/2026-data-streaming-report/#:~:text=investment%20in%20AI%20itself,production%2C%20with%20the%20majority%20experiencing)).
- 95% - Proportion of organisations reporting no measurable return on their AI investments to date ([www.sei.cmu.edu](https://www.sei.cmu.edu/news/sei-and-accenture-release-ai-adoption-maturity-model-to-help-organizations-scale-ai-with-predictable-outcomes/#:~:text=and%20build%20a%20clear%20roadmap,Mismatched%20expectations%2C%20misaligned%20applications%2C%20and)).

KEY TAKEAWAY

The week's developments underscore that success in AI now depends on enterprise data fundamentals. To turn AI hype into real value, leaders must invest in quality data, modern infrastructure, and governance - or risk falling behind competitors with stronger data foundations.

Sources

Confluent 2026 Data Streaming Report: AI Initiatives Stalled by Data Challenges
https://www.confluent.io/press-release/2026-data-streaming-report/

Accenture and the Carnegie Mellon University SEI Launch AI Adoption Maturity Model to Help Organizations Scale AI with Predictable Outcomes
https://newsroom.accenture.com/news/2026/accenture-and-the-carnegie-mellon-university-software-engineering-institute-launch-ai-adoption-maturity-model-to-help-organizations-scale-ai-with-predictable-outcomes

Salesforce Eyes $2B Listen Labs Buy as Fin Deal Closes, Chasing AI Data Moat
https://www.techtimes.com/articles/327263/20260910/salesforce-eyes-2b-listen-labs-buy-fin-deal-closes-chasing-ai-data-moat.htm

Salesforce and Google Cloud Unify Infrastructure and Agents for One Connected AI Stack
https://www.googlecloudpresscorner.com/2026-09-15-Salesforce-and-Google-Cloud-Unify-Infrastructure-and-Agents-for-One-Connected-AI-Stack

AI Governance in 2026: Is Your Organization Ready? – Dataversity
https://www.dataversity.net/articles/ai-governance-in-2026-is-your-organization-ready/