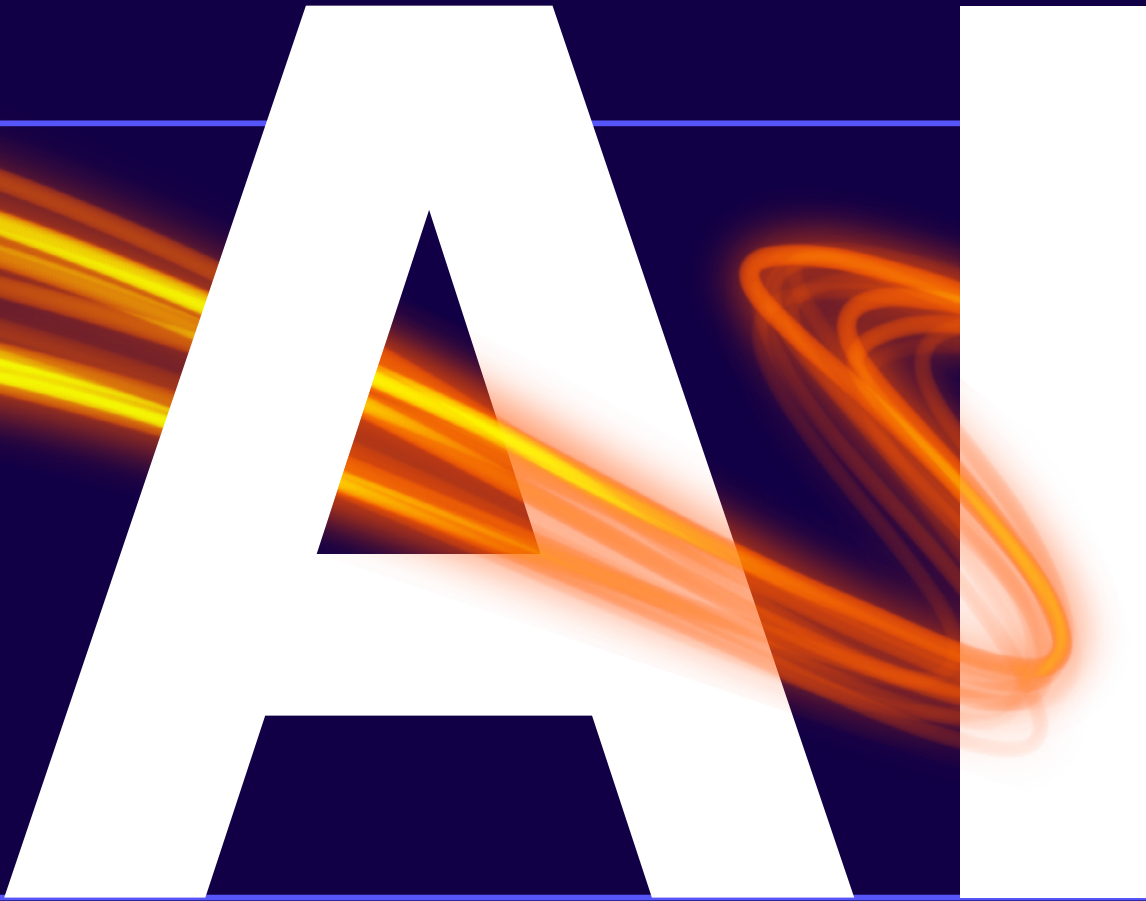


CLOUDERA



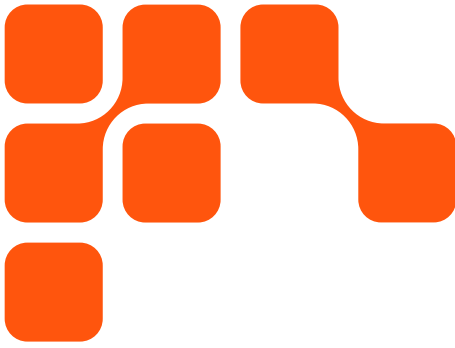
SURVEY REPORT

The State of Enterprise AI 2026



Table of Contents

Executive Summary	3
Introduction	4
Enterprise AI Meets Its Operational Era	4
The Economics of AI: Tokenomics, Cost and ROI	5
When AI Production Meets Reality	6
Governance Becomes the Price of Scale	7
Sovereignty Makes Hybrid a Business Requirement	8
If Enterprises Could Start Over	9
The State of AI Across Industries	10
What's Next for Enterprise AI?	11
Methodology	11
About Cloudera	12



Executive Summary

Enterprise AI continues to expand, but growing reliance on the technology is exposing new pressures around cost and control. Cloudera's 2026 State of Enterprise AI survey reveals five findings shaping the future of AI at scale:

63%

of respondents say AI is an important input in organizational decision-making, including 27% who say it regularly shapes important decisions or serves as a primary driver.

89%

of respondents have limited or redesigned AI use cases due to rising token consumption, showing how AI economics are already influencing enterprise strategy.

98%

of respondents have pulled back an AI deployment because of failure, hallucinations, or unexpected behavior, with 68% experiencing this sometimes or frequently.

75%

of respondents cite governance and compliance as a constraint on scaling AI, while 55% report faster-than-expected growth in governance and compliance costs.

58%

of respondents say moving AI workloads between environments requires moderate or major redesign, proving the difficulty of adapting deployments as requirements change.

Introduction

Last year, [Cloudera's 2025 State of Enterprise AI survey](#) captured a turning point for enterprise AI, with 96% of enterprise leaders reporting that AI had been integrated into their core business processes.

One year later, the question is no longer simply how enterprises adopt AI, but how they sustain and scale it as its role in the business grows.

Cloudera's 2026 State of Enterprise AI survey, conducted among 1,500 IT leaders and decision-makers at the director level or above across nine markets in the Americas, EMEA and APAC, examines that next stage of the enterprise AI journey. The findings show organizations navigating a more complicated operating reality as rising costs, governance requirements, production risk and changing workload requirements reshape the conditions for scale.

Perhaps nowhere is that shift clearer than in the economics of AI. Nearly nine in 10 respondents (89%) say their organization has limited or redesigned AI use cases because of rising token consumption. At the same time, governance is adding another layer of pressure: 55% say governance and compliance costs have grown faster than expected, while 75% identify governance and compliance as a constraint on scaling AI.

Token consumption

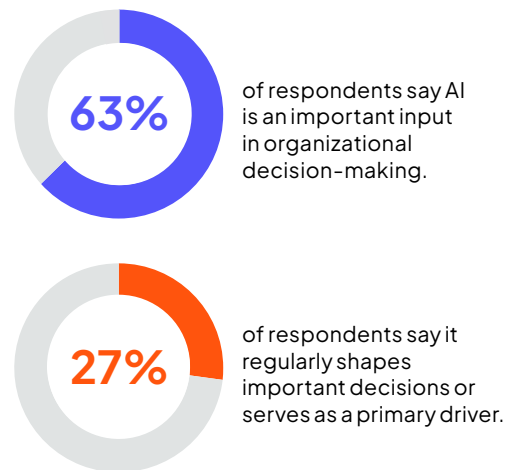
refers to the volume of information AI models process and generate, which directly influences the cost of running AI applications.

Enterprise AI is growing up, but maturity brings new tradeoffs. As organizations put more AI into production and entrust it with a greater role in the business, they must balance the drive to expand AI with the cost, control and flexibility required to sustain it. The challenge for enterprise leaders is increasingly not whether they can scale AI, but how they can preserve its velocity while adapting to the economic and operational realities that come with scale.

Enterprise AI Enters Its Operational Era

Enterprise appetite for AI remains strong, and its role in the business is deepening. AI is becoming a meaningful part of how organizations make decisions, with its influence extending well beyond experimentation and isolated use cases.

AI in organizational decision-making



Leaders are looking to extend that role even further. When asked about their most important AI priorities over the next 12 months, nearly half (43%) point to expanding AI adoption across the business, making it the most frequently selected response. APAC shows some indication of deeper reliance on AI at the more consequential end of decision-making, with 32% saying it regularly shapes important decisions or serves as a primary driver, compared with 27% overall.

The scale of investment behind it matches that influence. Nearly two-thirds (64%) of respondents say their organizations spend at least \$500,000 annually on AI initiatives, while 28% spend \$1 million or more. For many enterprises, AI now commands significant resources and plays a growing role in business operations and decision-making. With that deeper reliance comes a new set of considerations. Organizations have more at stake in how reliably and sustainably AI systems operate, extending the focus beyond getting AI into production to supporting it once it gets there.

Greater reliance on AI is putting the cost of sustaining that growth under greater scrutiny.

The Economics of AI: Tokenomics, Cost and ROI

AI's price tag is starting to shape its path through the enterprise. Last year's Cloudera survey found that cost was the leading barrier to AI implementation, while concerns about compute capacity had risen sharply from 8% in 2024 to 42% in 2025. This year, the impact is more tangible because rising costs put greater pressure on enterprises to demonstrate returns and manage the ongoing cost of AI at scale.

More than half (55%) of respondents say their AI investment has increased, but greater spending has yet to translate into clear returns for many enterprises. Fifty-eight percent say the value their AI initiatives generate is approximately equal to what they cost, while just (23%) say AI currently generates more value than it costs. 19% say costs outweigh the value generated.

The picture is similarly mixed when it comes to business outcomes. More than a quarter (27%) report little or no measurable impact from their AI initiatives, while 41% have encountered higher-than-expected costs. For organizations investing heavily in AI, these findings raise questions about how quickly that spending will translate into meaningful returns.

Nearly all enterprises believe they have some visibility into their AI spending. Ninety percent say they can measure end-to-end AI costs at least somewhat accurately. Yet only 27% say they can do so very accurately, suggesting there is a considerable gap between having a general understanding of AI costs and a precise picture of where that spending goes.



89%

of respondents have limited or redesigned their AI use cases due to rising token consumption.

This becomes even clearer at the level of token and inference costs. Only 19% continuously monitor those costs, even as they are already affecting how AI is used. Eighty-nine percent of respondents say their organization has limited or redesigned AI use cases due to rising token consumption.



Only 19%

of respondents continuously monitor token and inference costs.

Some of that pressure is more pronounced in EMEA. Forty-four percent of respondents there say token usage and model inference costs grew faster than expected, compared with 42% in APAC and 39% in the Americas. Data storage shows a similar pattern, at 43% in EMEA, versus 42% in APAC and 38% in the Americas. The differences are modest, but they point to specific parts of the AI cost stack where EMEA organizations are somewhat more likely to report faster-than-expected growth.

Token costs are changing AI strategy faster than many organizations can keep up with. As usage increases, enterprises will need greater visibility into variable AI costs to understand when an application remains economically viable and when changes are needed.

And organizations are already making those changes. Every respondent reports experiencing higher or more volatile AI costs, and 99% say their organization has taken action in response. Forty-five percent have moved a workload to another cloud, private environment or on-premises platform. Another 40% have reduced processing frequency, model size, context size, or output volume, while 24% have paused or canceled an AI use case altogether.

AI cost pressure is already prompting action.

99% of respondents have taken action in response to higher or more volatile AI costs.

AI economics can shift faster than enterprise plans, putting a premium on flexibility as usage grows. The fact that 45% have already moved workloads in response to cost pressure shows how directly AI economics can influence architecture. Workload placement may increasingly need to be treated as a decision that can change over time, making the ability to adapt without rebuilding applications from the ground up an important part of turning growing AI investment into sustainable returns.

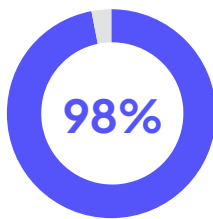
When AI Production Meets Reality

Putting AI into production gives enterprises a clearer view of where it delivers and where it falls short. Nearly all organizations (98%) have had to pull back an AI deployment at least once due to failure, hallucinations, or unexpected behavior. For more than two-thirds (68%) of respondents, pullbacks occur with some regularity, with respondents saying they have had to do so sometimes or frequently. Production can expose weaknesses that are harder to see during development or evaluation, making the ability to identify problems and adjust deployments an increasingly important part of operating AI at scale.

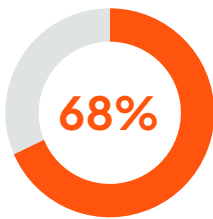
However, those setbacks tell only part of the story. Forty-three percent report greater-than-expected productivity, nearly twice the 23% who saw lower-than-expected gains. Widespread deployment pullbacks are therefore occurring alongside meaningful upside, complicating a simple success-or-failure narrative around enterprise AI.

Production experience reveals that AI can deliver value while still requiring organizations to adjust and reassess deployments. Managing that variability is becoming part of operating AI at scale.

AI into production



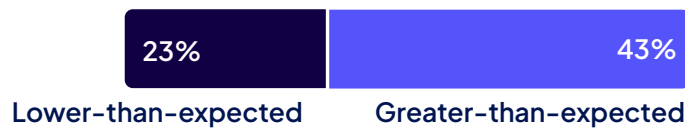
of respondents have pulled back an AI deployment at least once due to failures, hallucinations, or unexpected behavior.



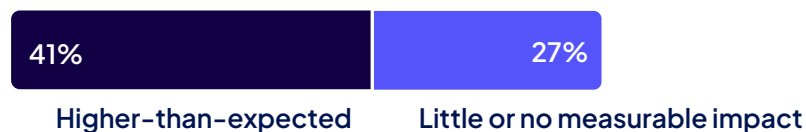
of respondents and almost one-third said they move data between environments daily.

Productivity is outperforming expectations for many enterprises, even as costs prove harder to predict.

Productivity



Cost Impact



Governance Becomes the Price of Scale

Governance is both a requirement for scaling AI and a source of friction in getting there. Three-quarters (75%) of respondents cite governance and compliance as constraints on scaling AI, making them the most frequently identified barriers. That control also comes at a growing price, with 55% saying governance and compliance costs have grown faster than expected, the highest of any AI cost measured.

Yet enterprises express considerable confidence in their ability to implement those controls. More than three-quarters (76%) say it is easy to apply consistent data governance, security and access policies across their AI environments. The disconnect suggests that implementing governance may be relatively manageable in isolation, while sustaining and extending it across a growing AI footprint introduces a much heavier operational burden.

That complexity is particularly evident in the Americas. Eighty-one percent of respondents there agree that AI integration makes data governance more complex and difficult to maintain, compared with 65% in EMEA and 61% in APAC. One possible explanation is that organizations in the Americas are managing AI across more established or expansive deployments, making the demands of maintaining consistent governance more apparent. The regional spread shows that the governance difficulties associated with AI integration are being experienced differently.

There is no single issue driving that pressure. Data privacy and third-party or vendor AI risk top the list of governance concerns, each cited by 35% of respondents. Data quality follows closely at 34%, with regulatory compliance at 33% and security and sovereignty at 31%. The breadth of those concerns explains why confidence in applying controls can coexist with governance remaining such a significant constraint on scale.

The challenge is finding enough control to support expansion without unnecessarily slowing it. Just over half (54%) believe their organization has struck an effective balance between the speed of AI adoption and appropriate governance. The rest are split evenly: 23% believe their organization is moving too cautiously, while another 23% say AI adoption is outpacing their governance and controls. That tension is pushing enterprises toward governance that responds to the level of risk involved rather than applying the same degree of oversight everywhere.

Seventy-eight percent require human approval for at least some AI-driven decisions, but the circumstances vary. More than half (52%) trigger approval when a confidence, risk or value threshold is crossed, while 34% require it for legally regulated or designated high-impact decisions and 30% for decisions affecting customers, employees or regulated outcomes.

Enterprises are drawing the line between AI autonomy and human intervention in different ways.

52%

approval triggered by risk or confidence thresholds.

44%

sampling or after-the-fact monitoring.

34%

approval for regulated or high-impact decisions.

30%

approval for decisions affecting people or regulated outcomes.

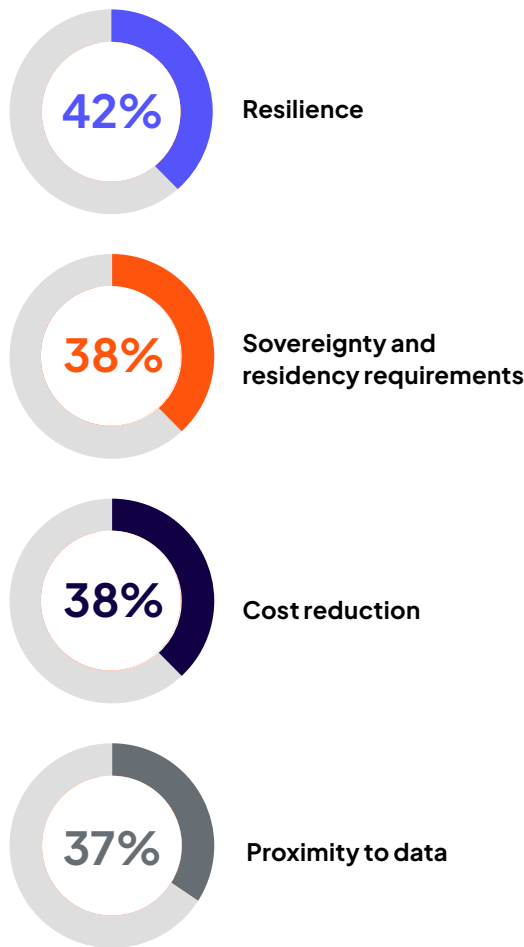
Approval is only part of how enterprises keep people involved. Forty-four percent also review AI outputs through sampling or after-the-fact monitoring. Human oversight is becoming more targeted, with organizations determining where AI can operate independently and where human judgment remains necessary.

Sovereignty Makes Hybrid a Business Requirement

The right environment for an AI workload today may look different tomorrow. Changing business requirements and regulatory obligations are making flexibility across environments increasingly important.

Enterprise data introduces geographic boundaries into an otherwise global AI landscape. Eighty-four percent of respondents say data residency, sovereignty or jurisdiction-specific requirements influence AI workload placement at least occasionally. For 39%, these requirements come into play frequently or always. The influence is somewhat more pronounced in the Americas, where 87% report considering these requirements at least occasionally, compared with 82% in EMEA and 79% in APAC. For global enterprises, AI strategy must account for the local requirements of the data that those systems depend on.

Why organizations operate AI workloads across multiple environments.



Most respondents operate across multiple environments, and the reasons reflect a range of business needs. Organizations are choosing where to run AI workloads based on the requirements of each use case, moving beyond a single deployment model. Resilience leads the considerations, while sovereignty, cost and proximity to data carry similar weight, reinforcing the importance of flexibility as requirements evolve.

But needing multiple environments does not mean enterprises can move freely among them. More than half (58%) say relocating an AI workload requires moderate or major redesign. At the same time, a sizable minority report less extensive redesign, showing that portability is not equally difficult for every enterprise. Moving the underlying data creates additional friction, with security risk and higher costs tied as the most frequently cited concerns at 34% each.

Cloudera’s 2026 Great AI Re-Architecture research found that 95% of enterprises had delayed AI projects due to infrastructure limitations. Against that backdrop, the gap between operating across environments and moving easily between them becomes even more notable. The findings from both surveys show that multi-environment AI is only as flexible as an enterprise’s ability to respond to changing workload requirements. Greater consistency in governance, security and infrastructure can reduce the work required to make those changes.

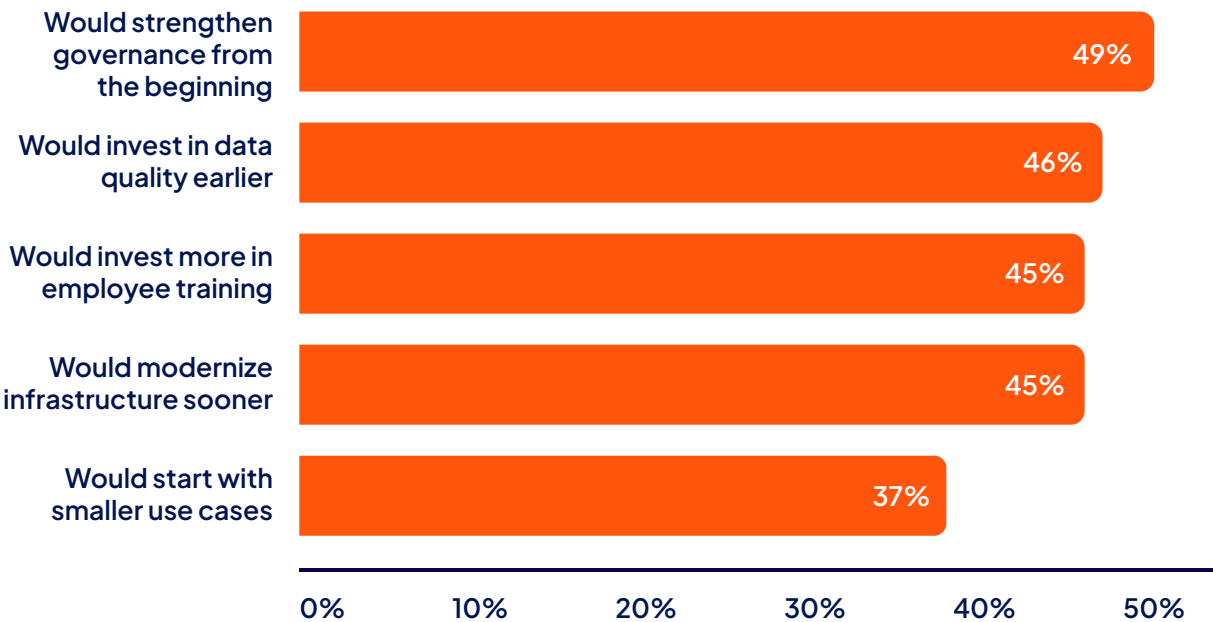
These limitations are also shaping how enterprises look back on their AI strategies, particularly the foundational investments they wish they had made sooner.

If Enterprises Could Start Over

Asked what they would do differently if they could restart their AI journey, respondents put the foundations first. Nearly half (49%) would strengthen governance from the beginning, 46% would invest in data quality earlier, and 45% would modernize infrastructure sooner. An equal share would invest more in employee training, while 37% would start with smaller use cases.

Experience is changing the order of operations for enterprise AI. Hindsight is telling: many of the investments enterprises prioritize today are ones they wish they had made earlier.

Many of the capabilities enterprises now see as essential to scale are the same ones they wish they had prioritized earlier.



The State of AI Across Industries

AI maturity looks different across industries, with each sector finding its own balance between cost, control and flexibility. For some, progress is reflected in stronger returns and greater confidence in AI-driven decisions. Others are navigating more demanding governance requirements or adapting their approach as production experience reveals where existing systems need to change. These differences reflect how industry-specific priorities shape what it takes to sustain AI at scale.

Pharma offers one of the clearest examples of AI taking on a more consequential role. Nearly half (49%) say AI regularly shapes important decisions or serves as a primary driver, compared with 27% overall. Financial services stand out for a different reason: 38% say AI generates more value than it costs, the highest share among the industries surveyed, while 70% have increased AI investment.

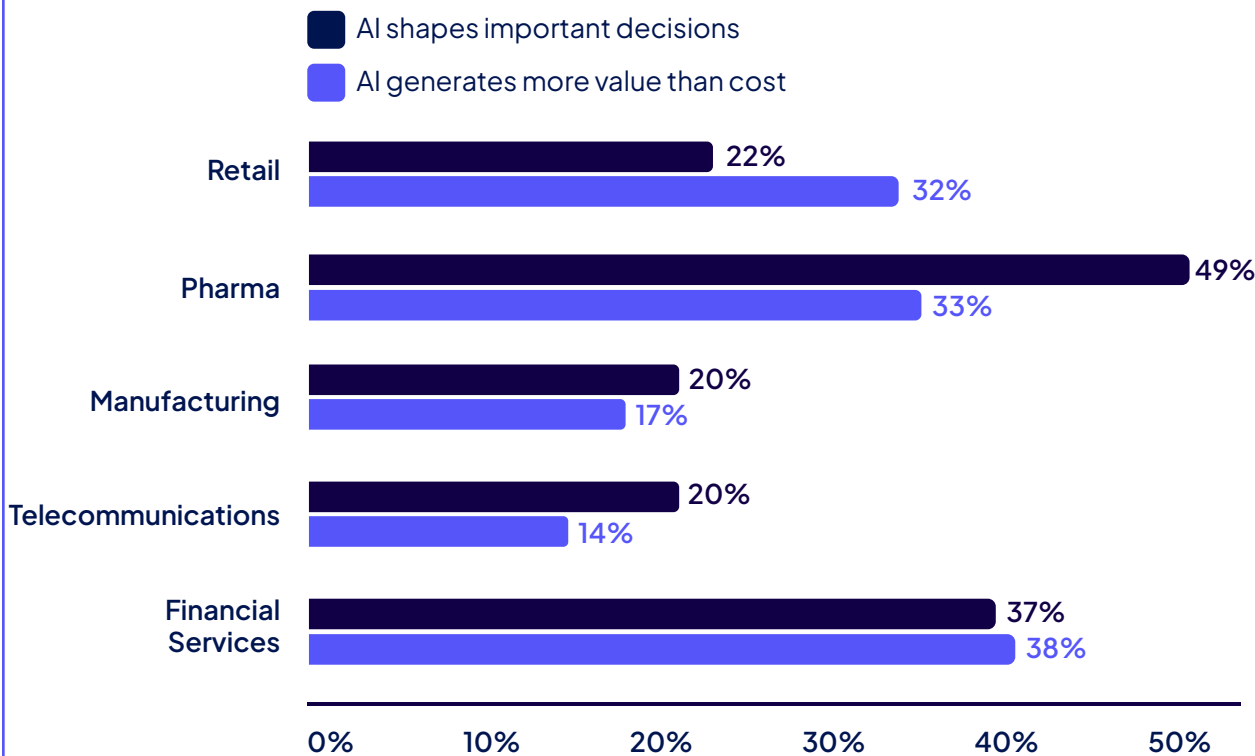
Manufacturing shows a more complicated picture. Nearly half (49%) report greater-than-expected productivity, but an equal share report higher-than-expected costs. Governance pressure is also particularly pronounced, with 84% citing governance and compliance as constraints on scaling AI.

Telecommunications exhibits another form of maturity altogether. Seventy-seven percent have sometimes or frequently pulled back an AI deployment after it reached production, yet 63% still say they have achieved an effective balance between AI speed and governance. That suggests frequent intervention may simply be part of operating AI at scale.

The industry cuts show that AI maturity takes different forms depending on the operating environment. Successful scale will depend on how well organizations adapt their AI strategies to the realities of their industries.

How AI Maturity Differs by Industry

Industries vary in both the responsibility they give AI and the economic value they report



What's Next for Enterprise AI?

What enterprises have learned from AI is to put adaptability at the center of what comes next. Models will change, economics will shift, and workload requirements will evolve. Enterprises need the freedom to respond while maintaining consistent control over the data and AI systems that support the business.

That makes flexibility a foundational requirement for AI at scale. Organizations need the flexibility to adopt new capabilities and change where workloads run without repeatedly rebuilding the controls and foundations that support them. The goal is to preserve the pace of innovation while ensuring consistent control wherever AI and data need to operate.

An enterprise data and AI platform can provide that foundation by bringing AI to data wherever it resides and applying consistent governance and control across environments. Cloudera enables enterprises to manage data and AI across their entire data estate, giving organizations the flexibility to respond as technology and business requirements evolve without sacrificing control.

Enterprise AI will continue to change, making the ability to adapt without losing control increasingly important. Building for that reality now can give enterprises more room to take advantage of what AI makes possible next. [Learn how Cloudera helps organizations](#) bring AI to their data wherever it resides while maintaining consistent governance and control across environments.

Methodology

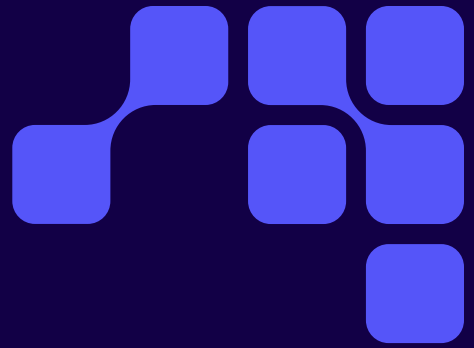
The survey, commissioned by Cloudera and fielded by Wakefield Research, examines the views of 1,500 IT leaders and decision-makers with a minimum seniority of director. The online survey was conducted in September 2026 across nine markets in three regions: the Americas (U.S., Canada and Brazil), EMEA (Italy, Germany and the UAE), and APAC (South Korea, Australia and Indonesia).



About Cloudera

Cloudera is the only hybrid data and AI platform company that large organizations trust to bring AI to their data anywhere it lives. Unlike other providers, Cloudera delivers a consistent cloud experience that converges public clouds, on-prem data centers, and the edge, leveraging a proven open-source foundation. As the pioneer in big data, Cloudera empowers businesses to apply AI and assert control over 100% of their data, in all forms, improving security, governance, and real-time and predictive insights. The world's largest brands across all industries rely on Cloudera to transform decision-making and ultimately boost bottom lines, safeguard against threats, and save lives.

To learn more, visit [Cloudera.com](https://cloudera.com) and follow us on [LinkedIn](#) and [X](#).



CLUDERA

Cloudera, Inc. | 6220 America Center Dr, 5th Fl, San Jose, CA 95002 USA | cloudera.com

© 2026 Cloudera, Inc. All rights reserved. Cloudera and associated marks are trademarks or registered trademarks of Cloudera, Inc. All other company and product names may be trademarks of their respective owners. 9978-01 October 16, 2026