January 2026
White Paper
Sponsored by Google Cloud

The AI Efficiency Gap:

From Total Cost of Ownership Crisis to Optimized Cost and Performance



Dave McCarthy
Research Vice President,Cloud and Edge Services, Worldwide Infrastructure Research, IDC



Warren de Lima
Research Manager, IDC



Table of contents



Click any title to navigate directly to that page.

IDC opinion 3

In this white paper 5

Situation overview 6

Measuring success 11

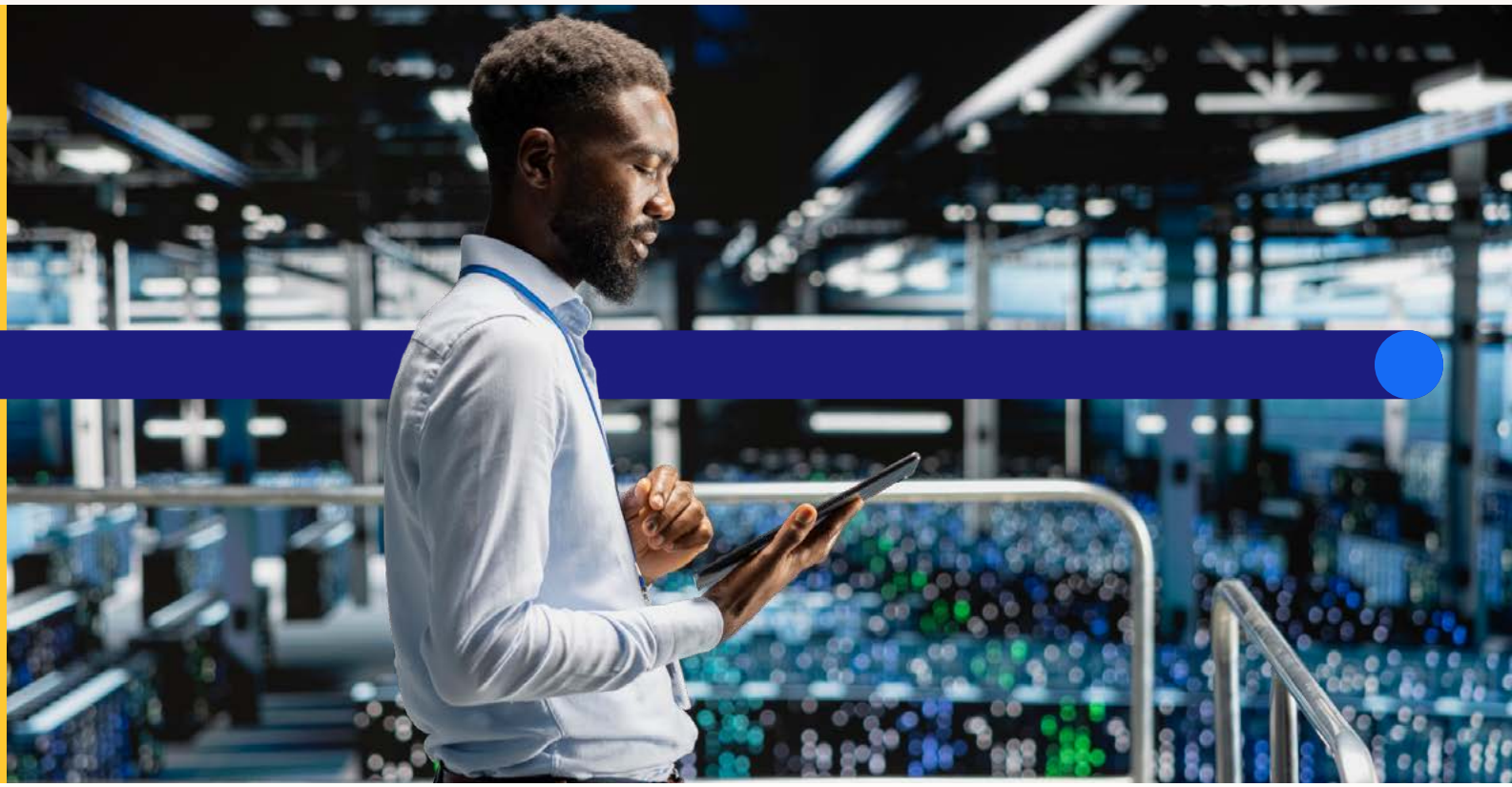
The impact of AI inefficiencies 15

Future outlook 21

Conclusion 22

Appendix: Accessible data tables 24

About the IDC analysts 25



IDC opinion

The rise of generative AI and the large-scale models that power them represent a fundamental architectural challenge to traditional cloud infrastructure. While general-purpose cloud workloads prioritize scalability and multi-tenancy, AI workloads introduce new, extreme demands that expose significant inefficiencies. This shift is forcing a complete reevaluation of infrastructure, moving the definition of success away from mere access to resources and toward total system efficiency.

These new workloads are architecturally different from legacy applications. Training large models requires massive clusters of specialized accelerators, such as GPUs and TPUs, which must work in parallel and communicate constantly. This, in turn, demands an extremely high-bandwidth, low-latency "fabric" that is far more robust than standard cloud Ethernet because the primary bottleneck is often the network itself.

Additionally, petabyte-scale data sets feed these models, requiring a high-performance storage system that can stream data at incredible speeds to prevent the expensive accelerators from sitting idle. This infrastructure must also serve two distinct purposes: the long-running, monolithic jobs of AI training and the short, real-time, bursty requests of AI inference.

Attempting to run these demanding workloads on legacy or general-purpose cloud infrastructure creates a significant "AI efficiency gap" — the difference between the theoretical performance of the AI technology stack and the actual performance achieved. A host of inefficiencies that result in a full-blown total cost of ownership (TCO) crisis drive this gap.

Organizations face high costs from the expensive specialized accelerators, and the problem is compounded because an end-to-end AI workload requires far more than just accelerators. Key components, such as data preprocessing and front-end user experiences, run primarily on standard compute. Inference engines frequently connect to systems of record (such as enterprise databases), which also rely on standard compute infrastructure.

When this entire environment — comprising specialized compute, standard compute, networking, and storage — is stitched together from different, non-optimized services, the resulting fragmented systems introduce significant operational complexity. Simply managing this complexity becomes a primary blocker to success, and a single bottleneck in one component can bring the entire GPU cluster to a halt.

A key element of this TCO crisis is the rampant underutilization of resources. When a GPU is idle — waiting for data from storage or for an update from another GPU over a slow network — it is still powered on and fully paid for. A 30% utilization rate on an accelerator cluster doesn't just mean a project is slow; it means 70 cents of every dollar spent on that infrastructure is wasted. This efficiency gap represents a massive financial drain that makes AI initiatives economically unsustainable at scale.

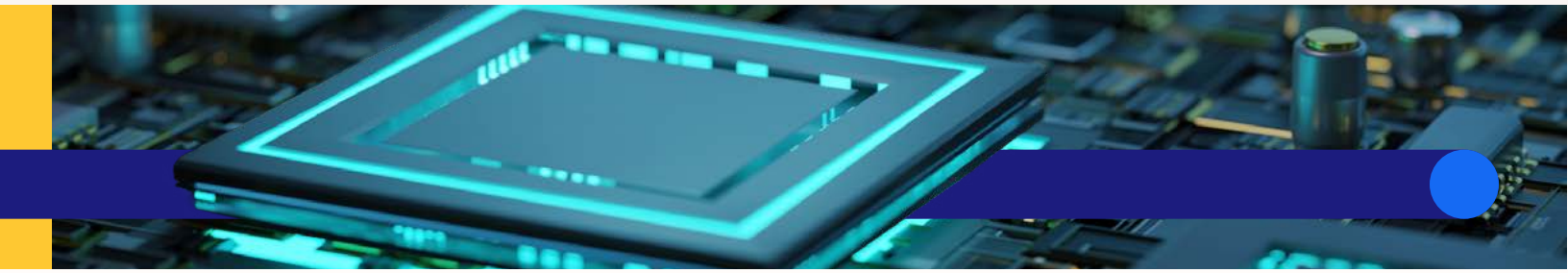
This economic reality fundamentally changes how organizations must define and measure the success of their AI platforms. The old metrics of simple availability, VM provisioning speed, or cost per VM hour are no longer sufficient. The new metrics are accelerator utilization, total time to train, total cost to train a model, performance per dollar for inference, and performance per watt. Success is no longer just about gaining access to 1,000 accelerators; it's about the economic viability of using those 1,000 accelerators effectively.

The path forward is to adopt a new model of AI infrastructure that directly addresses this TCO crisis. This model must be integrated with the entire stack of compute, networking, storage, and software; designed and optimized as a single, cohesive system to eliminate bottlenecks. It must be system-aware, meaning the software and orchestration layers understand the underlying hardware topology to intelligently place workloads, maximize data locality, and increase utilization.

Finally, this approach must be cost-optimized. By using an integrated and aware stack, organizations can dramatically increase resource utilization, which is the only sustainable way to manage the massive TCO of AI. The future of AI infrastructure is not about buying more powerful parts but about building more intelligent, holistic systems that waste nothing.

In this white paper

IDC conducted a global survey of 1,317 senior AI decision-makers to explore trends in AI infrastructure, including architecture decisions, business benefits, technical challenges, and how organizations are measuring success. This IDC white paper analyzes the results of this survey and presents a path forward for maximizing investments in AI infrastructure.



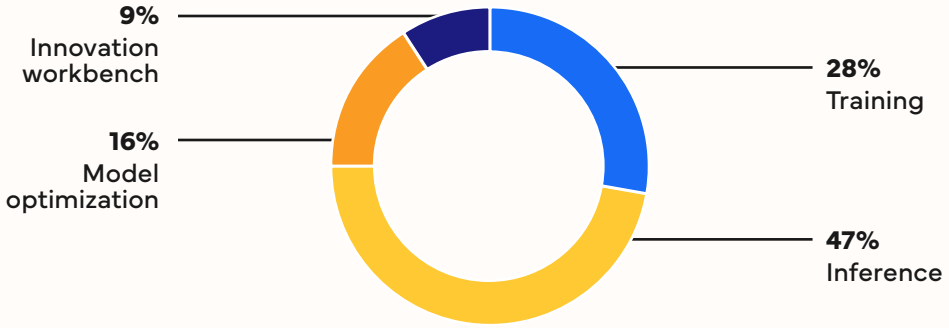
Situation overview

AI workloads can be thought of in four phases (Figure 1, below).

Figure 1

Four phases of AI workloads

Thinking about all your AI workloads, what percentage of them are in the following training, inference, model optimization, and innovation workbench phases?



n = 1,317; Source: IDC's *Machine Learning Efficiency Trends Survey,* August 2025

Definitions of the four phases of AI workloads:

- **Training:** The process of teaching a model by feeding it massive data sets, requiring large-scale, tightly coupled clusters of accelerators running for extended periods.
- **Inference:** The process of using a trained model to make predictions on new, live data, characterized by real-time, low-latency requests.
- **Model optimization:** The phase that includes fine-tuning, quantization, and distillation to make a trained model more efficient for deployment, balancing performance, cost, and accuracy.
- **Innovation workbench:** The experimental phase where data scientists and researchers explore new architectures, test hypotheses, and prototype models, requiring flexible and on-demand access to resources.

Survey respondents reported that inference has become the largest AI workload. This is because it represents the continuous, high-volume application of trained models to deliver real-time business value. While training is an extremely compute-intensive, expensive, and usually periodic event, inference is the non-stop, scaled-out phase that executes every time an AI application is used.

The massive scale of real-world usage drives the dominance of inference; a single model, once trained, serves billions of user queries, where every chatbot response or recommendation is a separate inference event. The sheer aggregate volume of these small, frequent requests ultimately dwarfs the total compute hours needed for training. The pervasiveness of AI in production amplifies this demand, especially with the growth of generative AI, where executing a full LLM response involves processing many tokens, leading to longer and more complex inference workloads per query, all requiring real-time, low-latency performance.

Most survey respondents (Figure 2, next page) confirm that cloud infrastructure is the preferred option for AI workloads. Cloud can uniquely provide:

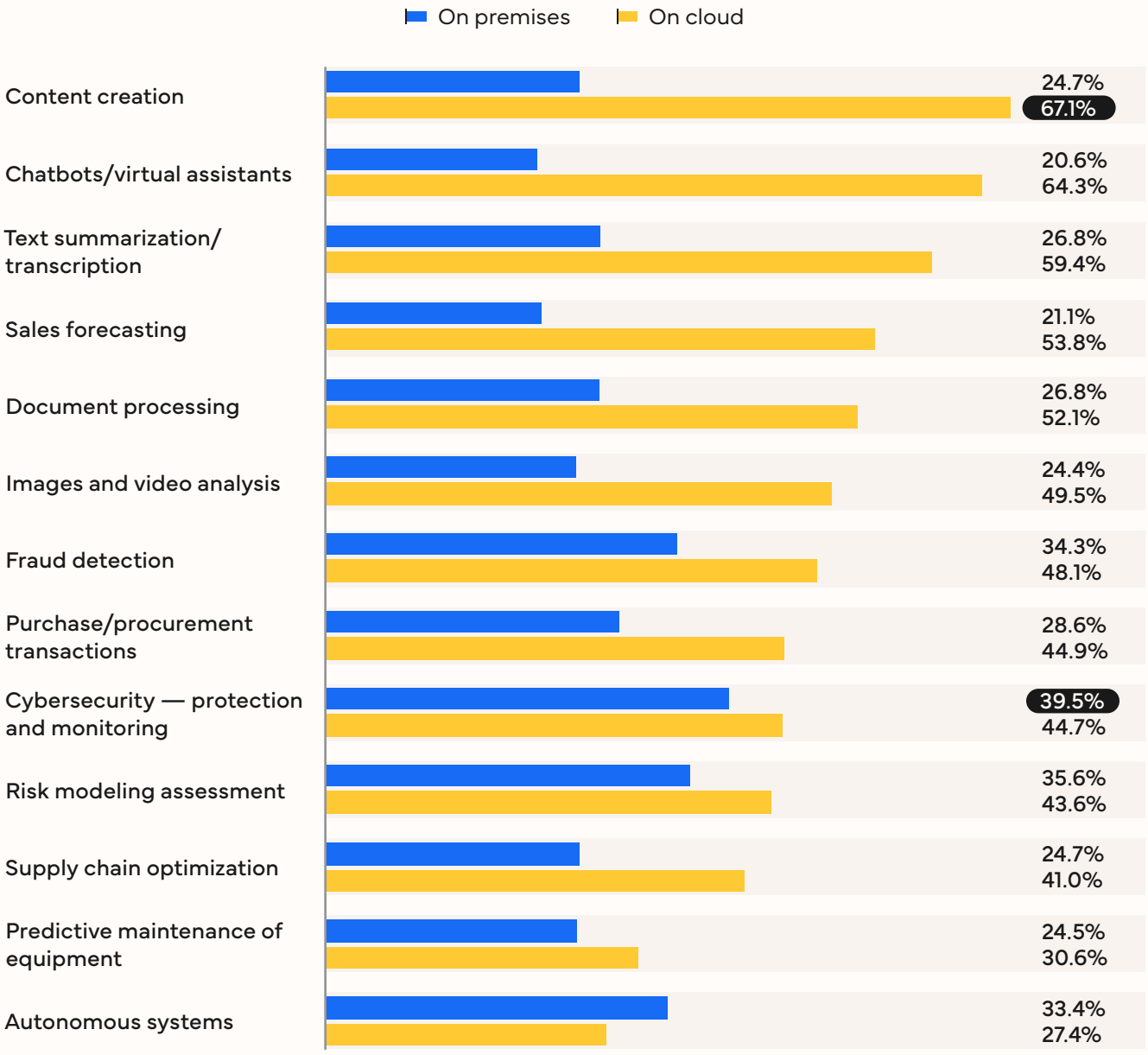
- **On-demand access to specialized hardware:** Cloud infrastructure offers on-demand access to specialized hardware, which is essential for modern AI workloads. Instead of incurring massive capital expenditures to purchase, cool, and maintain powerful GPUs and TPUs, development teams can provision these resources instantly, paying only for the compute time they use. This capability allows researchers to rapidly scale up to large clusters for intensive training jobs and then scale back down to zero, making experimentation and development cycles both faster and vastly more cost-effective than using static, on-premises datacenters.
- **Massive elasticity and scalability:** AI development requires massive elasticity and scalability that only the cloud can reliably deliver. When training cutting-edge models, compute demand can surge, requiring hundreds of interconnected processors for weeks; the cloud can provision this temporary scale instantly. Once models are deployed for real-time use (inference), cloud platforms utilize auto-scaling groups to dynamically adjust resources based on user traffic spikes, ensuring low latency and high availability globally without the need for constant, manual infrastructure management.
- **Adjacency to AI models and development tools:** Cloud providers ensure adjacency to AI models and integrated development tools, creating an efficient, end-to-end MLOps ecosystem. These platforms offer managed services that handle the complexities of data labeling, model tracking, and pipeline orchestration, abstracting away infrastructure management.

Figure 2

AI workload placement

Which of the following AI or ML workloads are deployed on premises or on the cloud?

See the figure data in an accessible table format.



n = 1,317; Source: IDC's *Machine Learning Efficiency Trends Survey,* August 2025



Although cloud is the preferred platform for most AI workloads, there are situations where deploying AI on premises or at the edge is desirable.

For applications dealing with real-world, physical systems, moving the AI processing closer to the data source can help address requirements related to:

- **Latency-sensitive applications:** Latency is a critical factor for real-time AI workloads such as robotics and predictive maintenance. Autonomous systems must react in milliseconds to ensure safety, prevent equipment damage, or maintain production quality. Sending massive sensor data (e.g., from an industrial robot or machine sensor) to the distant cloud for processing and waiting for the command to return introduces unacceptable network delays. Deploying the AI model at the edge ensures sub-millisecond decision-making, enabling instant anomaly detection, safe path correction, and immediate machine shutdown when needed, which is impossible to guarantee with central cloud infrastructure.

- **Regulatory compliance and data sovereignty:** For many organizations, data residency laws or internal security policies mandate that sensitive operational data cannot leave the company's physical premises. Industries such as defense, finance, and certain manufacturing sectors often treat operational logs, proprietary process data, and video feeds as highly restricted intellectual property. Deploying AI on premises or at the edge guarantees that this sensitive information remains within the secure, controlled environment and complies with regulatory requirements (such as the GDPR). Additionally, edge deployment allows the AI system to function autonomously even in remote locations with poor or nonexistent cloud connectivity.
- **Cost of data movement:** Edge devices generate enormous volumes of raw, continuous data (e.g., high-resolution video streams or high-frequency sensor readings). Transferring all this raw data to the cloud for processing is often prohibitively expensive due to high network bandwidth requirements and the cloud provider's data egress fees. By processing the AI inference locally at the edge, companies can filter and analyze the raw stream and only send back small, highly condensed insights or summary alerts to the cloud. This strategy drastically reduces the transferred data volume, which significantly cuts cloud transfer costs and saves bandwidth.

Cloud providers are addressing the need for on-premises and edge AI deployment by creating a hybrid cloud continuum that seamlessly extends their core services to customer locations. They offer dedicated, managed hardware appliances that are installed on premises but run the native cloud software, APIs, and services.

This provides a consistent environment for running AI workloads that require ultra-low latency or must comply with data residency regulations, essentially bringing a slice of the public cloud closer to the source of data.



Measuring Success

AI investments are delivering tangible, measurable success across a variety of metrics. Organizations that have embraced AI are now reporting significant, positive outcomes that are fundamentally reshaping their operations and market performance.

According to the survey results, the most immediate return has been in boosting the efficiency of their staff. A notable portion of organizations — 28.9% of respondents — experienced a substantial lift in their workforce productivity, reporting an impressive average improvement of 14.4%. This means that employees are spending less time on tedious, repetitive tasks and more on complex, strategic work.

This efficiency gain extends beyond individual workers to the entire operational landscape. Nearly a third of respondents, 31.8%, have seen the speed of their operational processes increase, reporting an average improvement of 10.1%. This acceleration, through automated workflows and optimized resource allocation, translates directly into faster time to market and enhanced production capacity.

28.9%
of respondents experienced a substantial lift in their workforce productivity.

AI is also elevating the quality of leadership and governance. Decision-making efficiency has soared, with 28.7% of respondents observing an average improvement of 11.7%. AI-powered analytics and forecasting provide leaders with rapid, comprehensive insights, allowing them to make faster, more confident, and data-backed choices that drive business forward.

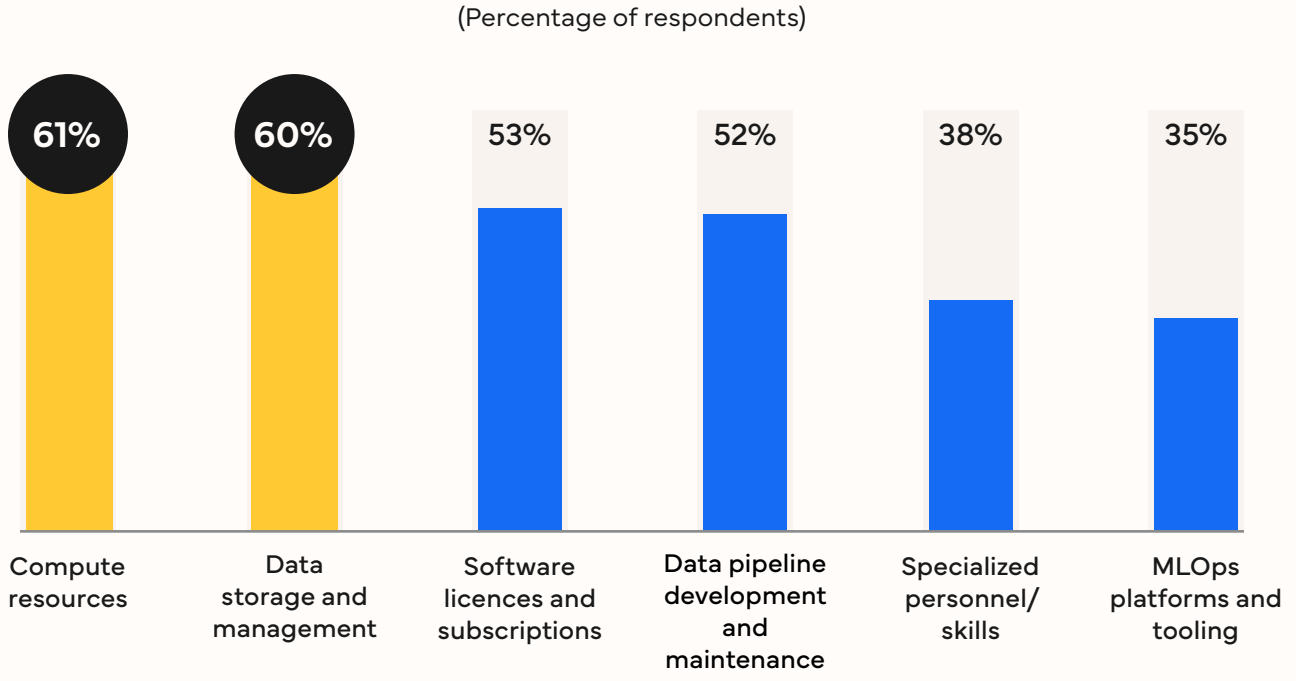
Ultimately, these internal improvements resonate powerfully with the end user. The most human-centric benefit is the improved customer satisfaction and experience. Over a quarter of organizations — 28.3% of respondents — reported an average improvement of 14.1% in this critical area. They often achieved this through AI-driven personalization, 24 x 7 responsiveness via intelligent agents, and predictive services that anticipate customer needs, solidifying loyalty and driving repeat business.

Today, most of the focus of evaluating the efficiency of AI workloads has been on understanding the impact of technology costs. **Figure 3 (below)** highlights that compute resources, data storage, and data management have the highest impact on TCO calculations. These are followed by software licenses and subscriptions and data pipeline development and maintenance.

Figure 3

Impact on total cost of ownership

Which of the following contributes the most to the TCO for your organization's AI initiatives?



n = 1,317; Source: IDC's *Machine Learning Efficiency Trends Survey,* August 2025

However, many organizations find it challenging to quantify the true value and effectiveness of their AI investments because success metrics often fall into two distinct, hard-to-reconcile categories: high-level business outcomes and detailed technical performance. Simply deploying a model isn't success; measuring success requires a structured approach to defining and tracking both types of metrics.

Business metrics focus on the strategic impact of AI, translating model performance into clear financial or operational benefits. Organizations struggle to connect model outputs directly to these large-scale results.

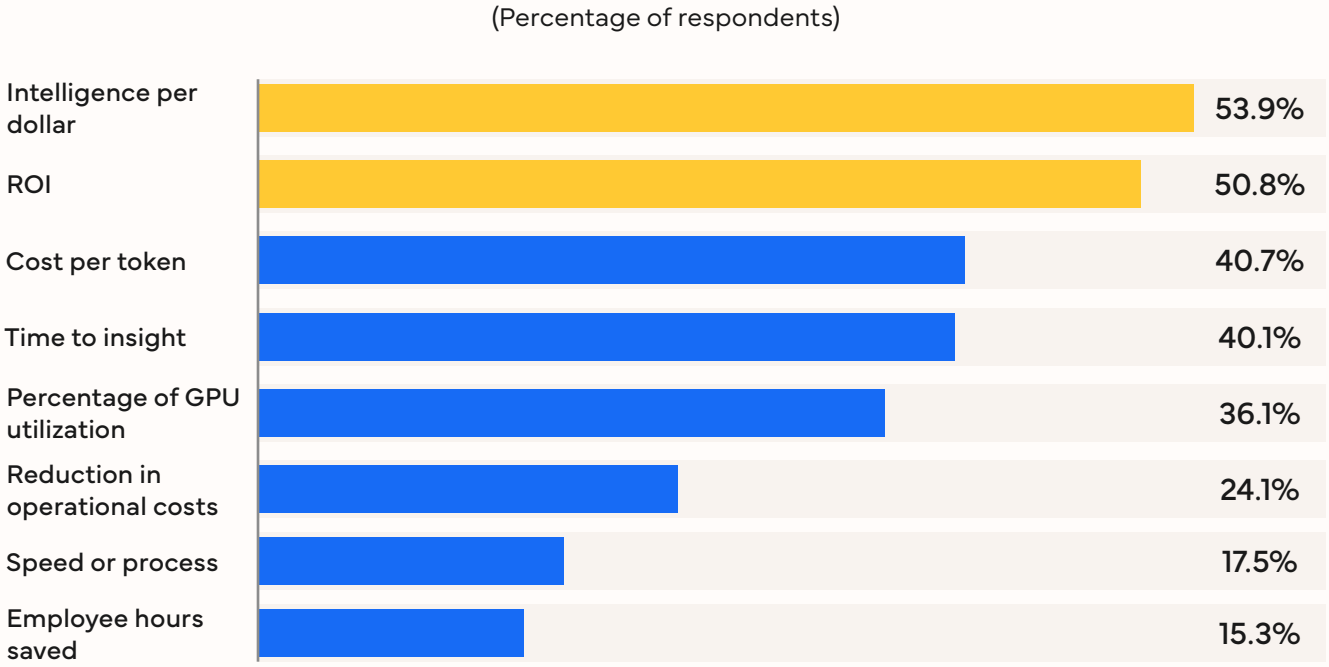
Technical metrics focus on the quality and performance of the AI model itself. While these are essential for data scientists, they often lack immediate meaning for business stakeholders, making it difficult to justify continued investment.

In many cases, challenges in measuring success are related to limitations in the AI stack. **Figure 4 (below)** represents the top metrics that organizations want to measure but are unable to do so due to these limitations.

Figure 4

Measuring AI success

What are three things you would like to measure but are unable to do so due to limitations in your AI stack?



n = 1,317; Source: IDC's *Machine Learning Efficiency Trends Survey,* August 2025

While not an industry standard, "intelligence per dollar" can be thought of as the total cost per unit of useful intelligence delivered to an end user. LMArena exemplifies this concept because its ranking system, combined with available pricing data, forces a direct comparison between a model's measured utility and its price.

A similar metric is "goodput," which is a measure of the useful or valuable output of an AI system, as opposed to its raw processing speed.

- In the context of large-scale model training (such as for LLMs), the metric is often called ML Productivity Goodput. It measures the true efficiency of a massive, long-running training job.
- In the context of AI inference (i.e., serving a model to users), goodput is a user-centric metric that combines performance with quality of service.



The impact of AI inefficiencies

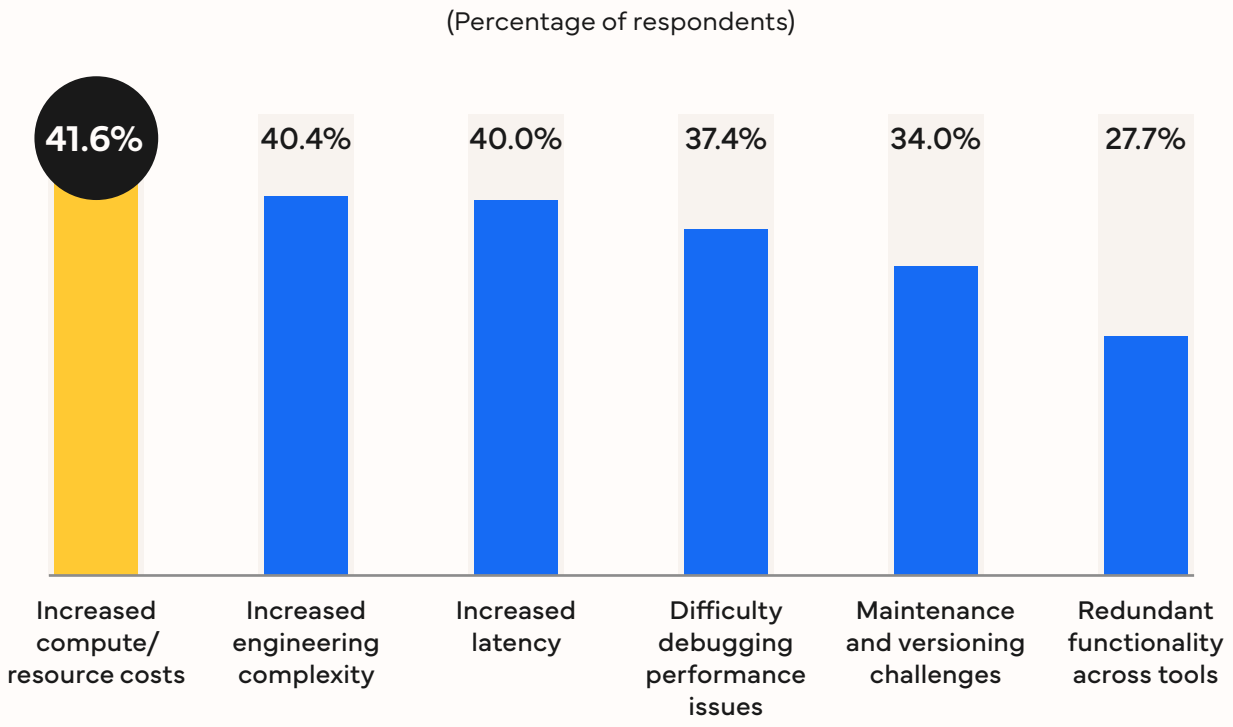
A significant majority of organizations, specifically 54.3% of survey respondents, frequently or sometimes utilize multiple AI frameworks (such as TensorFlow, PyTorch, or JAX) or multiple hardware platforms (such as GPUs from different vendors, specialized AI accelerators, or CPUs) across their infrastructure. This patchwork approach is often the result of different teams independently selecting tools based on historical projects or specific model requirements.

When this fragmentation occurs, the effect is overwhelmingly negative: 92% of organizations report a negative effect on efficiency. The impact of these inefficiencies is tied to increased costs, complexity, and performance **Figure 5 (next page)**.

Figure 5

Impacts on efficiency

Which of the three biggest inefficiencies has your organization experienced as a result of multiple AI frameworks or hardware platforms in a single workload?



n = 1,317; Source: IDC's *Machine Learning Efficiency Trends Survey,* August 2025

Every 1% gain in efficiency reduces TCO significantly because compute and storage are the most significant drivers of cost. Since large-scale AI services, such as real-time product recommendation engines, process billions of transactions annually, even a minor reduction in computational overhead yields enormous savings.

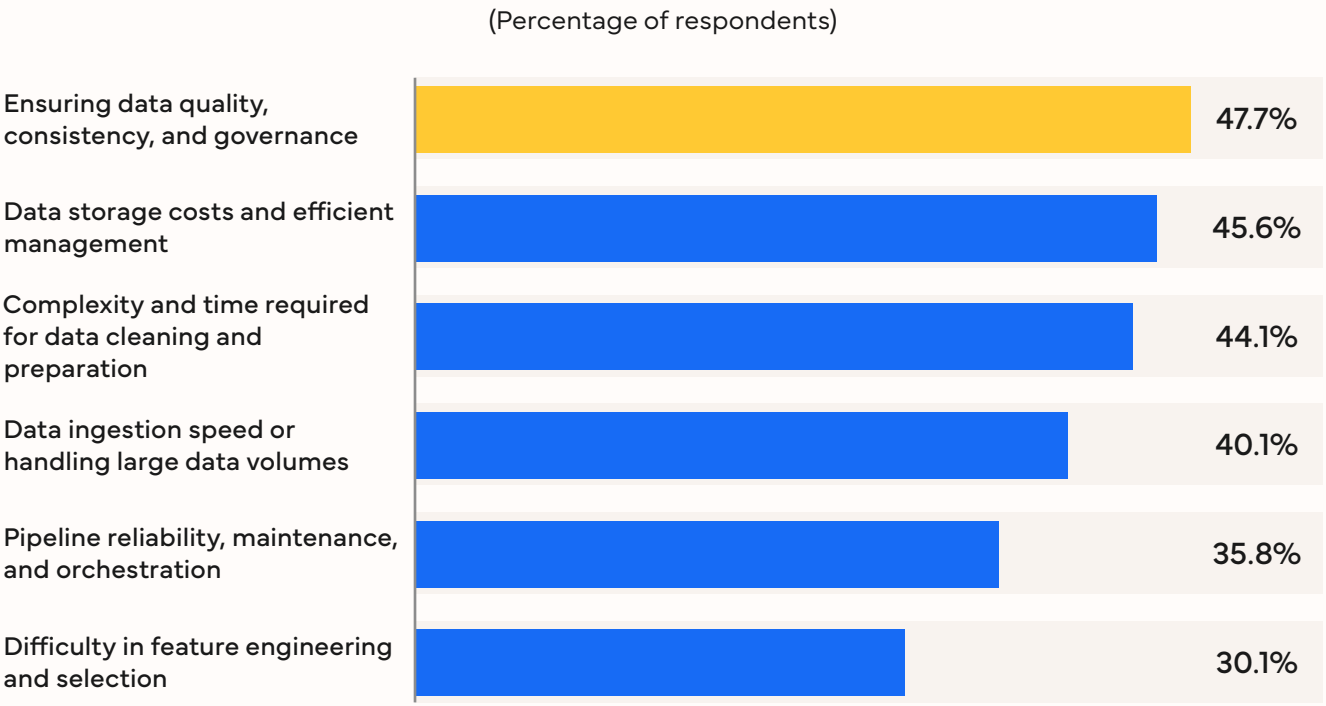
Consider an application generating 500 million inferences daily. If the annual compute cost is, for example, $182,500, a mere 1% efficiency gain (reducing the compute time or resources necessary per inference) directly results in an annual saving of over $1,800. This seemingly small reduction in TCO is persistent and scales across every AI model and operation within the organization. When this saving is multiplied across dozens of high-volume services and includes associated reductions in power consumption and cooling, a 1% technical improvement quickly translates into hundreds of thousands or millions of dollars in sustainable operational cost reduction.

Given the data-centric nature of AI, data management also contributes to the list of challenges. **Figure 6 (below)** illustrates the main challenges relating to managing data pipelines that affect business outcomes.

Figure 6

Data management challenges

Which, if any, of the following challenges are you currently experiencing within your AI/ML data pipelines that make it difficult to achieve your business outcomes?



n = 1,317; Source: IDC's *Machine Learning Efficiency Trends Survey*, August 2025

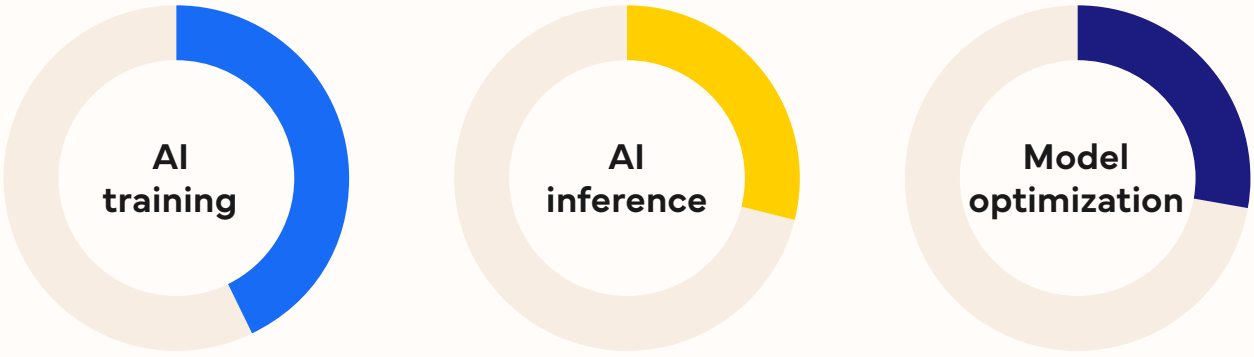
The impact of AI inefficiencies is motivating a closer look at where budget waste is happening across the different phases of AI. While training represents the largest amount of budget waste, there are impacts in all three areas **Figure 7 (next page)**.

Figure 7

AI budget waste

To the best of your knowledge, how much of your AI budget was spent on AI tools that did not deliver on expected ML productive value?

(Percentage of respondents)



n = 1,317; Source: IDC's *Machine Learning Efficiency Trends Survey,* August 2025

The most direct form of waste comes from expensive hardware sitting unused. Idle GPU time during experiments or deployment accounts for significant inefficiency:

- **Training:** During the resource-intensive model training phase, idle time often occurs between experiments, while engineers are preparing data, or due to poor job scheduling that leaves powerful GPUs or TPUs waiting.
- **Inference:** In production deployment, idle time can result from inconsistent request traffic. When an inference cluster is provisioned for peak demand, it will sit largely idle during low-traffic periods, wasting near-constant operational costs.

Some 25.7% of survey respondents indicated that idle GPU time impacted budget waste for training; 24.9% saw an impact on inference operations.

Inefficient use of resources represents waste that occurs even when the hardware is technically active. Inefficient use of resources indicates that the compute units are running, but the models or software are not optimized to fully utilize the available processing power:

- **Training:** This is the single largest identified area of waste. It happens when training jobs are poorly configured, code is unoptimized, or the chosen framework doesn't efficiently leverage the specific hardware architecture.
- **Inference:** In deployment, inefficiency can arise from a lack of model quantization, the use of inefficient data pipelines, or poor batching strategies, requiring more compute cycles per prediction than necessary.

Some 27.5% of survey respondents indicated that inefficient use of resources impacted budget waste for training; 22.3% saw an impact on inference operations.

Overprovisioned GPU/TPU clusters are the problem of building or allocating more infrastructure than is truly needed, often as a buffer against peak loads or due to poor capacity planning:

- **Training:** Teams frequently provision large clusters for training projects to ensure quick turnaround times, but those clusters often sit underutilized once the initial resource-heavy phase is complete.
- **Inference:** Overprovisioning in deployment is a common strategy to guarantee low latency and reliability during peak demand. However, maintaining this excess capacity for the majority of the time results in a substantial financial drain, driving up the total cost of ownership.

Some 24.2% of survey respondents indicated that overprovisioned clusters impacted budget waste for training; 21.9% saw an impact on inference operations.

To effectively address the issues relating to inefficiencies and significant waste, organizations are making key strategic investments. They place the highest priority on adopting cloud optimization tools, with 30.4% of organizations focusing their efforts here to better manage and scale their infrastructure costs **Figure 8 (next page)**.

27.5%
of survey respondents indicated that inefficient use of resources impacted budget waste for training.

To effectively address the issues relating to inefficiencies and significant waste, organizations are making key strategic investments.

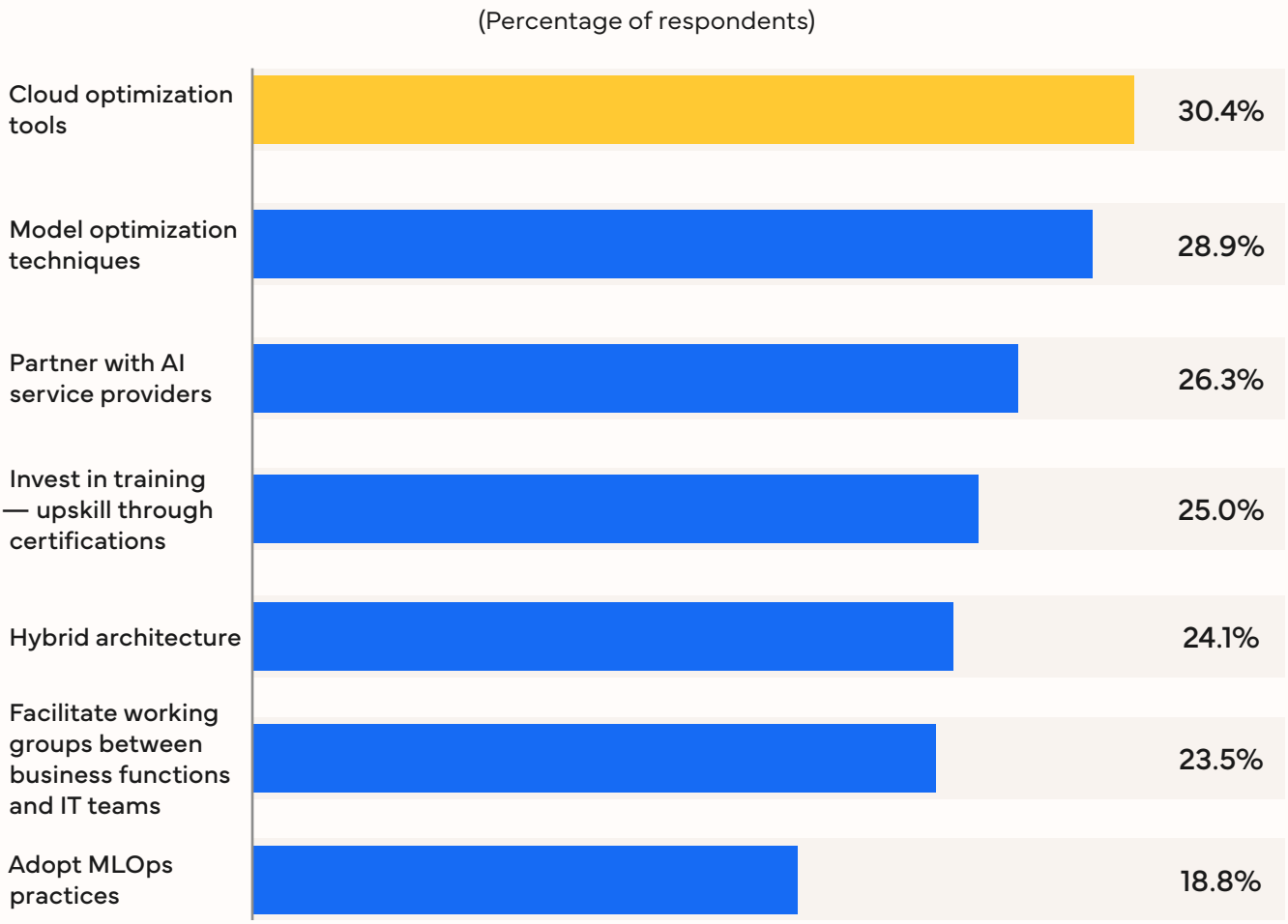
Closely following is the commitment to technical excellence: 28.9% are actively implementing model optimization techniques to ensure their AI runs as efficiently as possible, maximizing hardware utilization.

Finally, recognizing the complexity of these challenges, 26.3% of organizations are choosing to partner with specialized AI service providers to gain external expertise and access to best practices for efficiency and cost control.

Figure 8

AI investments

What initiatives is your organization prioritizing to overcome those challenges?



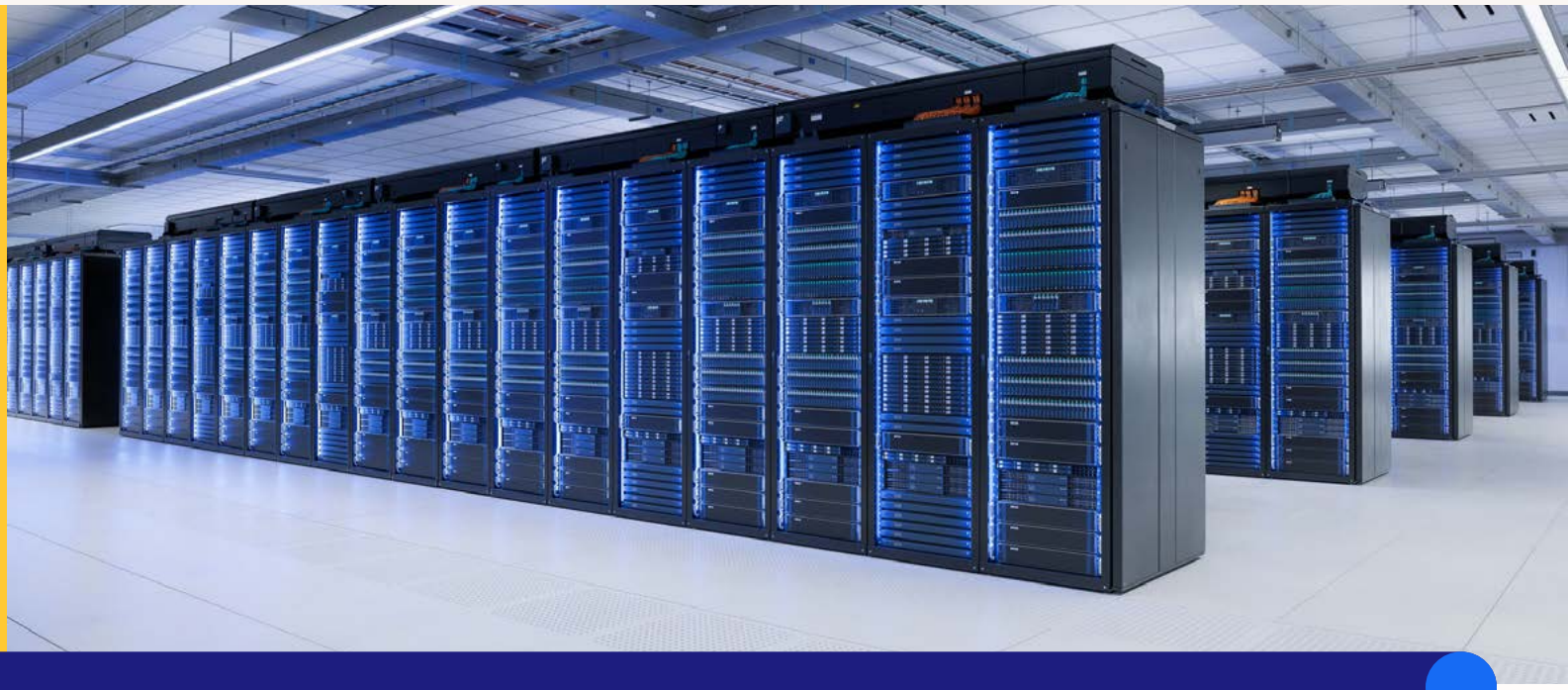
n = 1,317; Source: IDC's *Machine Learning Efficiency Trends Survey,* August 2025



Future Outlook

Organizations are looking ahead and clearly identifying three primary operational and business hurdles they expect to face with their AI initiatives over the next two years. These challenges center on costs, talent, and measurable value.

- **Controlling the rising costs of AI:** The top concern for organizations is managing the financial aspects of AI. Some 32.6% of respondents are worried about controlling and optimizing the rising costs of AI development and deployment. This concern reflects the high expense associated with procuring and maintaining specialized hardware, the continuous training of large models, and the expanding infrastructure footprint necessary to support massive AI inference loads in production environments.
- **The AL/ML talent shortage:** Closely following cost is the challenge of securing the necessary human capital. Some 31.5% of respondents anticipate difficulties with finding, hiring, and retaining skilled AI/ML talent. The demand for experienced data scientists, ML engineers, and MLOps specialists far outstrips the supply. This talent scarcity drives up salaries and makes it hard for organizations to scale their AI teams, slowing down the development and effective deployment of AI solutions.
- **Measuring and demonstrating clear ROI:** Many organizations struggle with proving that their AI investments are worthwhile. Some 29.8% of respondents predict issues with measuring and demonstrating clear ROI for AI investments. This reflects the difficulty in translating technical AI metrics (such as model accuracy) into concrete business outcomes (such as revenue increase or cost savings). If leaders cannot clearly quantify the financial benefit of an AI project, future funding and organizational buy-in for new initiatives become significantly harder to secure.



Conclusion

Optimize for TCO, not list price, to accelerate time to market: Organizations should shift their financial perspective from just comparing hardware or software list prices to calculating the TCO. The true TCO crisis is not simply about wasted budget; it represents the massive opportunity cost of delayed innovation and unrealized business value. To address this, organizations must move from upfront, static pricing models to a system of total cost visibility. This involves adopting dynamic workload scheduling, implementing consumption-based models for compute resources, and tracking system reliability metrics. It is also crucial to factor in hidden costs such as downtime, system fragmentation, latency, and the availability and cost of skilled talent.

Simplify the setup and management of the AI stack: System-level complexity is a key driver of inefficiency and developer friction. To address this, organizations must simplify the entire AI environment. This starts by consolidating frameworks and hardware under managed, interoperable systems to eliminate the waste that fragmentation causes. Using managed, system-aware software reduces the constant engineering overhead necessary to manually manage diverse components. Organizations should also prioritize portability and cross-accelerator compatibility. This prevents vendor lock-in and provides the necessary flexibility to quickly test new frameworks and hardware without needing to reengineer the entire solution.

Eliminate waste and bridge the boardroom gap with better measurement: Effective monitoring and measurement are essential to eliminate waste and demonstrate AI's value to leadership. Organizations must implement telemetry to proactively track metrics such as idle compute time and efficiency KPIs. Using real-time monitoring and automated scaling capabilities is critical to cutting underutilization and ensuring expensive hardware is always busy when needed. Success must be demonstrated by accounting for both technical KPIs (such as latency and utilization) and business KPIs (such as ROI and revenue lift) to prove the project's success aligns with the organization's overall strategy.

Modernize data infrastructure: The foundation of any efficient AI initiative is its data infrastructure. Organizations must proactively address the top data challenges, specifically data storage costs and ingestion speed. This requires moving beyond general-purpose storage and adopting AI-specific storage and networking solutions designed for the high-throughput, high-parallelism demands of modern AI training and inference. To simplify procurement and management, organizations should partner with providers offering a unified stack that tightly integrates AI models, data management tools, and the underlying IT infrastructure.

Closing the AI efficiency gap requires a holistic strategy that simultaneously addresses TCO, simplifies the fragmented AI stack, eliminates waste through rigorous measurement, and modernizes data infrastructure. These proactive steps are essential for converting raw AI potential into consistent business value.

Ultimately, success hinges on the strategic decision to partner with a provider offering a unified stack and the expertise to help optimize the cost and performance of AI initiatives, ensuring that every dollar invested yields maximum intelligence and sustained competitive advantage.

Appendix: Accessible data table

Figure 2 accessible data

AI workload placement

AI or ML workload	On premises	On cloud
Content creation	24.7%	67.1%
Chatbots/virtual assistants	20.6%	64.3%
Sales forecasting	26.8%	59.4%
Document processing	21.1%	53.8%
Images and video analysis	24.4%	49.5%
Fraud detection	34.3%	48.1%
Purchase/procurement transactions	28.6%	44.9%
Cybersecurity — protection and monitoring	39.5%	44.7%
Risk modeling assessment	35.6%	43.6%
Supply chain optimization	24.7%	41.0%
Predictive maintenance of equipment	24.5%	30.6%
Autonomous systems	33.4%	27.4%

n = 1,317; Source: IDC's Machine Learning Efficiency Trends Survey, August 2025

Return to figure

About the IDC Analysts



Dave McCarthy

Research Vice President,Cloud and Edge Services,
Worldwide Infrastructure Research, IDC

Dave McCarthy is research vice president within IDC's Worldwide Infrastructure Research organization and global research lead for the Cloud and Edge Services practice. McCarthy leads a team of analysts covering research on shared (public) cloud, dedicated (private) cloud, and edge deployments, services, adoption trends, vendor strategies, and market dynamics. Benefiting both technology suppliers and IT decision-makers, McCarthy's insights delve into ways in which hybrid and distributed cloud platforms provide the foundation for next-generation workloads, enabling organizations to innovate faster, automate operations, and achieve digital resiliency..

More about Dave McCarthy →



Warren de Lima

Research manager, IDC

Warren de Lima is a Research Manager at IDC. He designs, executes, and analyzes quantitative and qualitative primary research that drives effective business decisions. His areas of research expertise include the following: brand health, shopper insights, purchase behavior, segmentation, new product research, and in-depth interviews. Warren has 15 years of research experience in Canada, the UK, and the USA. Throughout his career, he has had a successful track record of helping clients' re-brand and introduce new products. Warren has a B.A. in Law from Carleton University and an M.A. in European Politics, Business, and Law from the University of Surrey (UK). He has an interest in behavioral science and looks for opportunities to integrate it into his research designs to capture subconscious business decision-making triggers.

More about Warren de Lima →

IDC Custom Solutions

IDC Custom Solutions produced this publication. The opinion, analysis, and research results presented herein are drawn from more detailed research and analysis that IDC independently conducted and published, unless specific vendor sponsorship is noted. IDC Custom Solutions makes IDC content available in a wide range of formats for distribution by various companies.

This IDC material is licensed for external use and in no way does the use or publication of IDC research indicate IDC's endorsement of the sponsor's or licensee's products or strategies.



idc.com @idc @idc

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications, and consumer technology markets. With more than 1,300 analysts worldwide, IDC offers global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries. IDC's analysis and insight helps IT professionals, business executives, and the investment community to make fact-based technology decisions and to achieve their key business objectives.

©2025 IDC. Reproduction is forbidden unless authorized. All rights reserved. CCPA