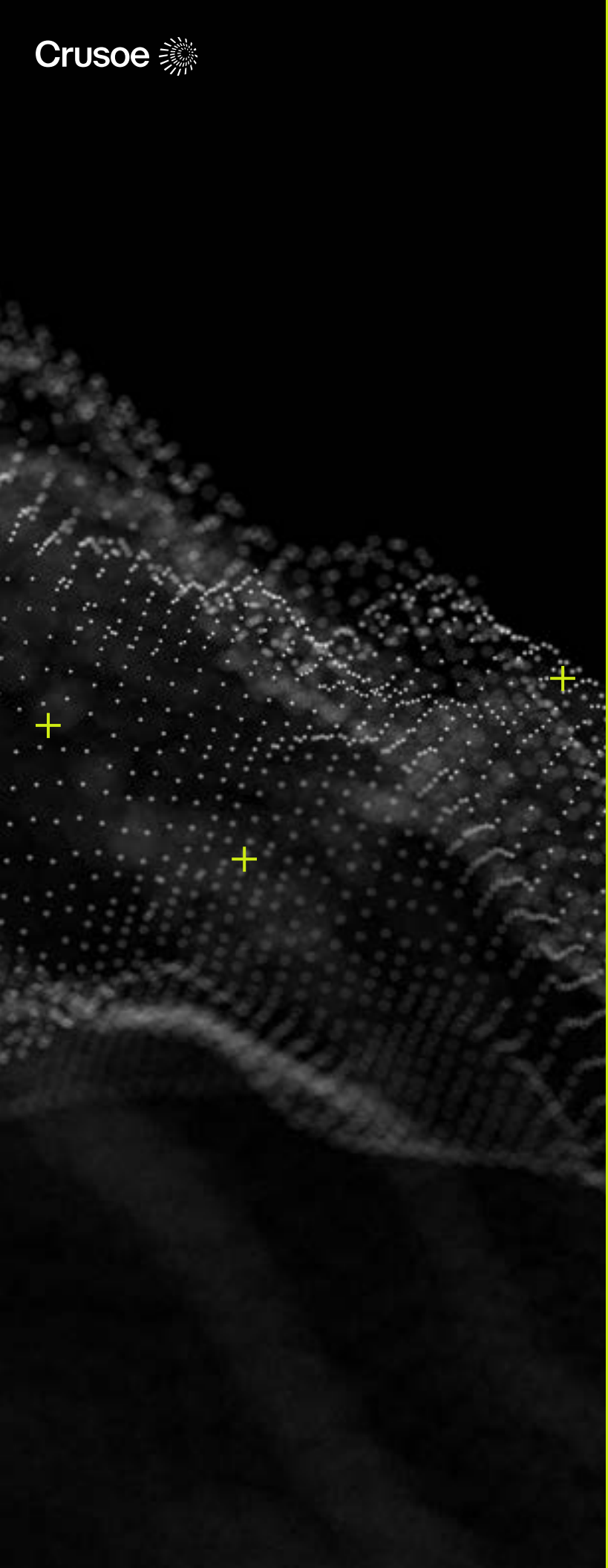
Crusoe



2026

# AI infrastructure trends report

# Contents

# Executive summary

Not long ago, AI was considered an experiment, handed off to tech teams eager to explore its possibilities. Today, it's rewiring how businesses operate, solve challenges, and drive growth. Across every industry, teams are racing to tap into the value of large language models (LLMs), automation, and next-generation intelligent tools as AI becomes an essential part of everyday work.

## The AI imperative

To better understand how organizations are navigating this shift as they move into 2026, Crusoe partnered with MetaLab to survey and interview enterprise and digital-native business (DNB) leaders on the front lines of ambitious AI imperatives.

The findings paint a clear picture: teams are turning to AI to clear away friction and free up resources — not simply to do more with less, but to create space for creativity and innovation to flourish. One respondent put it this way: "AI will be central to our innovation strategy, helping us stay ahead of industry trends and deliver cutting-edge solutions that differentiate us in the market."

But this race to realize AI's promise is far from straightforward. Despite strong ambition, most organizations face a host of infrastructure-related barriers.

## Barriers to building the future

The struggle persists because the traditional cloud ecosystem was not designed for the age of intelligent automation. Decision-makers in our research surface a common set of challenges:

- Performance bottlenecks and downtime that erode confidence and slow AI project delivery
- Unpredictable and opaque costs that make it nearly impossible to scale responsibly
- Fragmented support and a lack of domain expertise, leaving organizations with gaps between ambition and execution
- Limited control over critical infrastructure, making teams vulnerable

Throughout this report, we examine these persistent barriers — and, more importantly, explore strategies for overcoming them.

## Infrastructure as the new differentiator

Forward-looking teams are rethinking not just how they use AI, but also how they build it. They're finding a better way with infrastructure partners who can enable speed and performance without compromising security, sustainability, or ease of use. This is the shift Crusoe is pioneering: a vertically integrated approach, built from the ground up for AI. This AI factory model vertically integrates energy, hardware, data center, cloud, and managed AI services into a single, cohesive platform. It's a new paradigm that accelerates value creation while aligning with the demands of both today's business and the future well-being of people and the planet.

If you're navigating AI initiatives, our research offers a roadmap — and a call to action. We'll show how a new, purpose-built approach to AI infrastructure is enabling organizations to move faster, with greater control, to create measurable impact.

# Introduction

Across nearly every sector, the conversation around AI has changed. Rather than asking where AI can add value, teams are now actively seeking ways to translate AI's potential into real business results at scale and at speed. In practice, this means deploying AI for fraud detection, predictive analytics, intelligent document processing, supply chain optimization, customer service automation, personalized marketing, and more.

Implementing AI should be straightforward, but the reality for most organizations is riddled with roadblocks. Inconsistent access to high-performance compute, sluggish deployment cycles, operational complexity, limited control over data environments, and unpredictable costs all conspire to slow progress. The result is a widening gap between what's possible with AI and what's actually being delivered. We're at an inflection point as we look to the future of what AI will enable. To get there, it's simply not enough to advance incrementally.

This report provides a data-driven perspective from leaders at the center of AI transformation, offering a blueprint for achieving ambitious goals through infrastructure that's designed for the future of AI.

# Key insights

### 1. Efficiency is a catalyst for innovation

AI is being deployed first and foremost to drive operational efficiency. But for leading organizations, efficiency is just the launchpad. By clearing friction and freeing up resources, AI is paving the way for fresh creativity, faster innovation cycles, and stronger business impact.

### 2. Dependability, cost-effectiveness, and transparency are non-negotiable

Companies building AI expect their cloud providers to deliver reliable performance, predictable costs, and transparency. Yet, most find that current providers — especially hyperscalers — fall short. Persistent gaps here limit business value and stall AI progress.

### 3. Greater control provides greater confidence

Leaders are increasingly uneasy with over-reliance on traditional cloud vendors and opaque systems. They're demanding more direct control over their AI infrastructure to minimize surprises, reduce risk, and ensure their goals aren't limited by someone else's roadmap. The shift toward end-to-end integration reflects a new expectation: infrastructure partners should clear hurdles and empower teams to bring their vision to life.

### 4. Expert support fills knowledge gaps

AI adoption is outpacing internal expertise. Builders are looking to cloud partners for proactive, expert support that can bridge internal knowledge gaps and accelerate success. Many are willing to switch providers for better support, even when technology is comparable.

### 5. Adaptability and ease of use are requirements

The growing complexity of AI and cloud environments is a weight on both business and technical teams. Companies building AI want adaptable, easy-to-use solutions that reduce the burden of integration, support a range of workflows, and make it simple for teams to deliver results at scale.

### 6. Responsible AI and sustainability are rising priorities

While not always the first filter for selection, responsible AI practices and sustainability are gaining ground — particularly for digital-native businesses and as ESG commitments shape long-term priorities. Providers who demonstrate a commitment to responsible AI and sustainable operations are increasingly favored.

# Methodology

Quantitative survey March 2025

## 332

Participants including:

**152**

**Decision-makers**
(C-suite, VPs, Executive Directors) from large enterprises ($500M+ in revenue or $8B+ valuation)

**70**

**Influencers**
(Directors, Senior Managers in AI/ML/Engineering)

**79**

**Decision-makers**
from digital-native businesses

**31**

**Practitioners**
(Data Scientists, Developers, Engineers)

Crusoe engaged MetaLab to conduct an in-depth, mixed-methods research study that would capture both the breadth and depth of decision-maker experience across industries.

Respondents were drawn from a cross-section of industries, including technology, media and entertainment, manufacturing, automotive/robotics, enterprise SaaS, and general software. All participants had direct responsibility for evaluating or implementing AI and cloud infrastructure within their organizations.

**Qualitative interviews**

In addition to the survey, in-depth interviews were conducted with select decision-makers. These conversations provided richer context and firsthand perspective on provider selection, pain points, and emerging needs.

**Sample rigor and composition**

- The research included organizations from North America and key global markets.
- Industry quotas ensured balanced representation across major sectors.
- All decision-makers surveyed had direct strategic or budgetary influence over AI/cloud investments, ensuring executive-level perspective and actionable insight.

**Stakeholder and ecosystem lens**

To further ground the research in real-world delivery, we integrated input from 15 internal Crusoe stakeholders and layered in competitive and trend analysis across the AI/cloud ecosystem.

The result is a research-backed view into what decision-makers truly need from their AI infrastructure partners, what stands in their way, and where the next opportunities for value creation are developing.

# 1

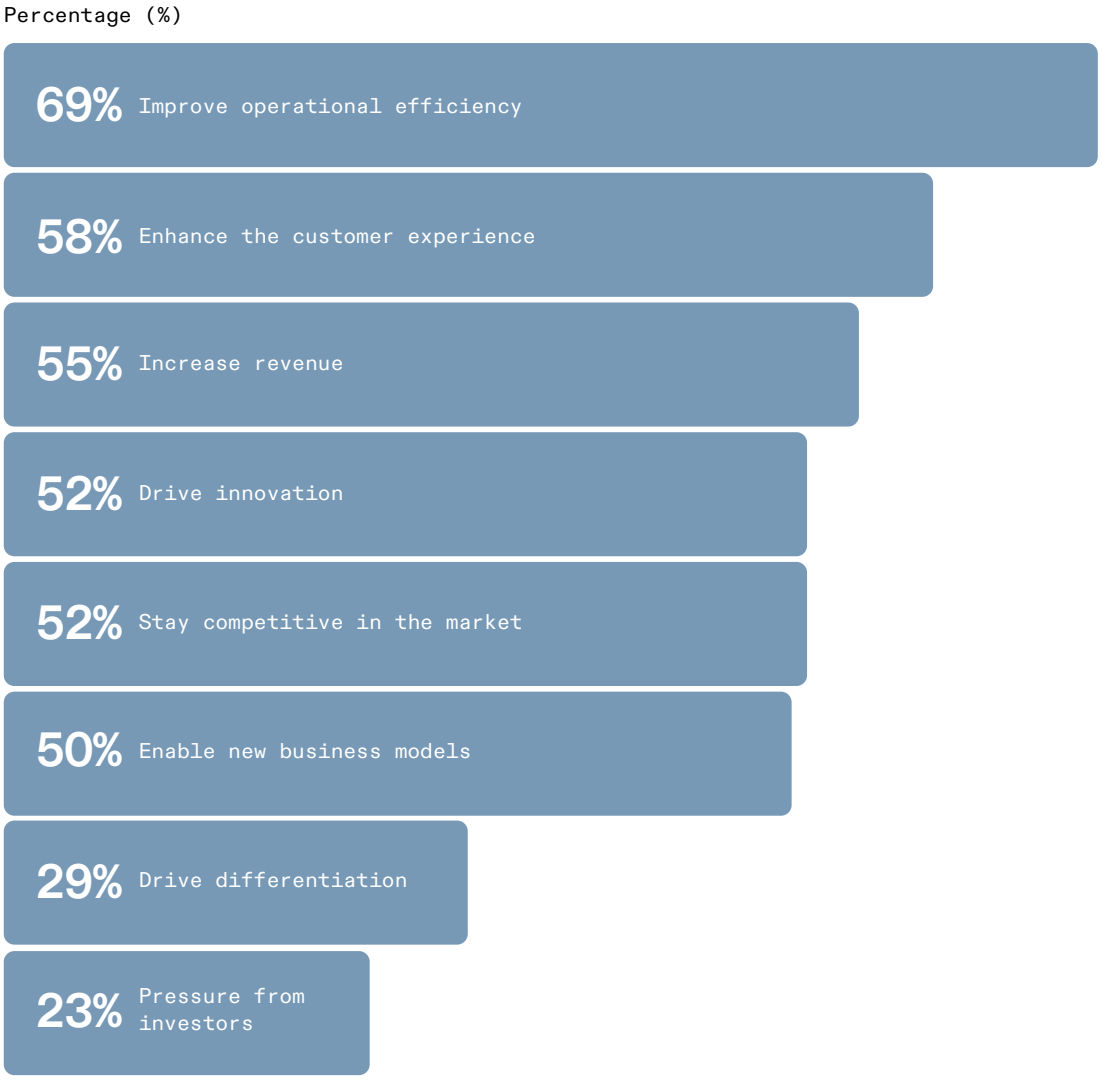
## Efficiency is a catalyst for innovation

For AI leaders, capturing the value of AI begins not with moonshots, but with immediate gains in efficiency. Across our research, the single most common motivator for AI investment is a mandate to streamline operations and free up resources. 69% of the leaders we surveyed say improving operational efficiency is their top priority for AI initiatives.

# Decision-makers' motivations for driving AI initiatives

But decision-makers are clear that operational efficiency is not the ultimate goal. Rather, it's a springboard for something greater. 55% of respondents prioritize AI as a means to increase revenue, and 52% see it as a pathway to drive innovation. This signals that efficiency gains are meant to fuel broader business value. 51% cite staying competitive as a primary motivator, recognizing that every advantage secured now could be the difference between leading competitors or playing catch-up.

This perspective is echoed across industries and organizational types in open-ended responses. One executive captured the sentiment this way: "By automating key aspects of our operations, we'll be able to drastically reduce inefficiencies and cut down on human error. This will free up our teams to focus on more creative and strategic tasks, leading to faster innovation cycles and better decision-making powered by data-driven insights."




## The bottom line:

Efficiency creates capacity for what comes next. Organizations that use AI to clear operational friction are positioning themselves to capture new opportunities, adapt to change, and lead their markets with speed and creativity.

# 2

## Dependability, cost-effectiveness, and transparency are non-negotiable

AI cloud infrastructure must do more than provide compute power. Our research reveals that today's leaders view cost control, security, performance, reliability, and scalability as table stakes for a successful AI initiative.

# AI cloud infrastructure must do more than provide compute power

Yet the reality of the current market often leaves organizations in the lurch. Many teams continue to encounter gaps where these expectations are concerned, especially as workloads scale in size and complexity. One leader put it this way: "The cloud provider we use across our enterprise must be easily maintained, be performant, be well supported, and meet our compliance standards. Otherwise, all our AI initiatives will just fail."

Our data reveals the pain points with hyperscalers around these non-negotiables:

- Security and compliance concerns were rated as the most pressing pain point, with an average relevance score of 6.86 out of 10.
- Performance issues (including downtime and latency) followed closely, at 6.84 out of 10.
- Scalability challenges — the ability to flex resources as business needs evolve — came third, at 6.65 out of 10.
- Cost management (unpredictable pricing, hidden fees) was rated fourth, at 6.58 out of 10.

| Attribute | Table stakes | Differentiated | Not important |
|---|---|---|---|
| Cost | ● | ○ | ○ |
| Security & compliance | ● | ○ | ○ |
| Performance | ● | ○ | ○ |
| Enterprise scale infrastructure | ● | ○ | ○ |
| Reliability | ● | ○ | ○ |
| Scalability | ● | ○ | ○ |
| Ease of integration | ● | ○ | ○ |
| Managed services (e.g. training inference) | ● | ○ | ○ |

Note: Table stakes criteria were consistent across organization type (enterprise vs. DNB) and industry

# Decision-makers' hyper scaler pain points

In other words, what many providers advertise as differentiators are actually the bare minimum needed for enterprises and DNBs to succeed with AI. A decision-maker explained: "One of the biggest challenges we face with [our cloud provider] is managing costs and scalability, as prices can escalate quickly with increased usage. This can limit our ability to scale AI initiatives efficiently and may require additional budget planning.


## The bottom line:

Uncompromising dependability, cost-effectiveness, and transparency are the new baseline for building with AI.

# 03

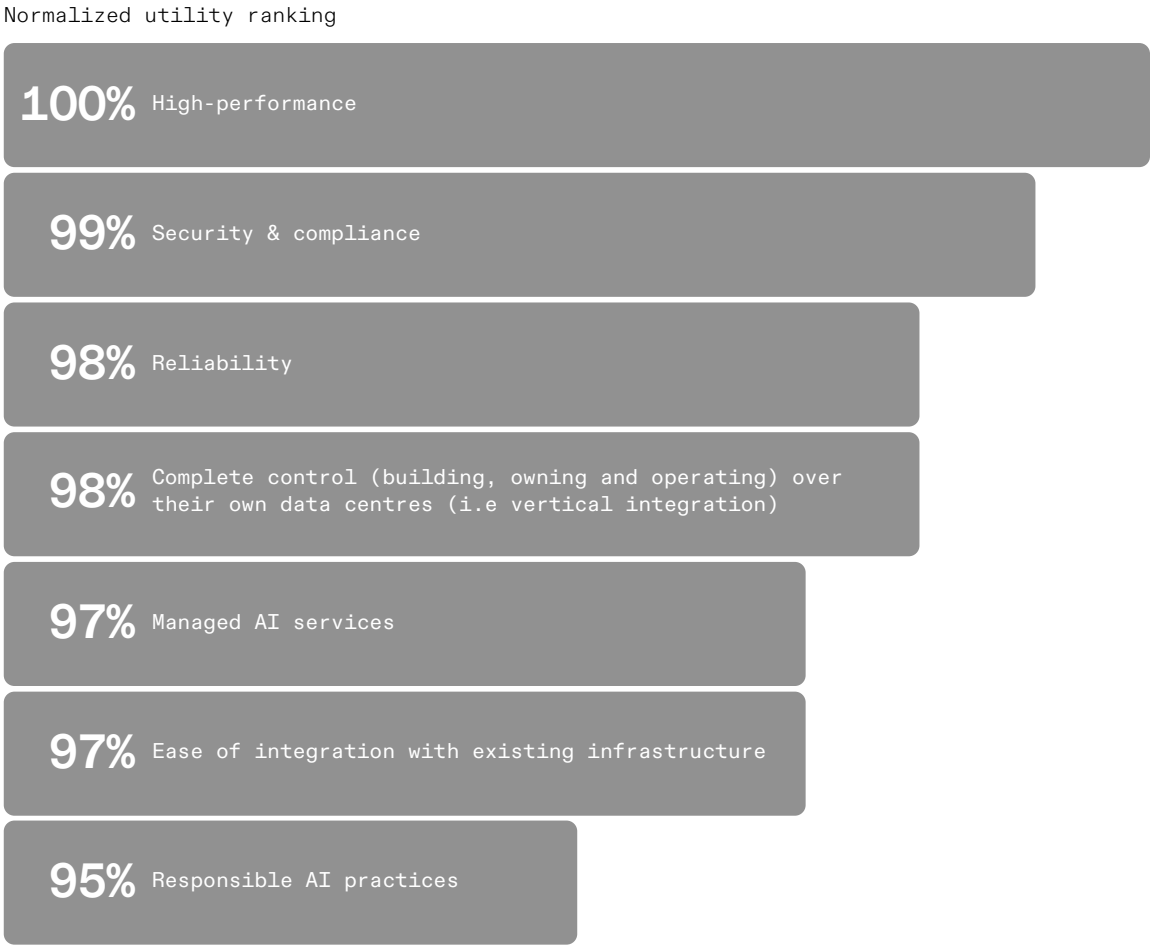
# Greater control provides greater confidence

As AI ambitions intensify, so do the anxieties around dependence on traditional cloud providers. Many leaders in our study voiced a growing sense of vulnerability as over-reliance on third-party clouds leaves organizations exposed to unexpected disruptions, cost spikes, performance inconsistencies, and shifting provider priorities.

This quest for agency is producing a shift in how decision-makers evaluate AI infrastructure partners. Leaders pinpoint the need for direct oversight of performance and security, rather than outsourcing core operations to fragmented supply chains or external vendors.

Vertical integration in an AI factory model, giving teams complete control over the full AI stack, is emerging as a highly differentiated attribute. In our study, 98% of decision-makers rated “complete control (building, owning, and operating) over their own data centers” as important. Notably, it was the only attribute more often seen as a differentiator than a mere table stake, outscoring even established priorities like performance, security, and reliability.

# Importance of cloud provider attributes for decision makers




One term that surfaced as especially resonant in the research is "the AI factory." While hyperscalers were designed for generic data storage and compute, decision-makers now expect providers to be "purposefully built for AI" — from the physical infrastructure to the software layer and everything in between.

**Defining the AI factory: A vertically integrated approach**

Legacy clouds were not built to deliver what today's leaders demand: independent control, transparency, and purpose-built performance. Enter the AI factory, a new paradigm for the future of enterprise AI.

The AI factory is an end-to-end, vertically integrated platform that unites every layer of the stack — from energy sourcing, to high-performance data center construction, to the AI cloud and managed services that bring it all together. Unlike legacy providers that patch together generic compute and storage solutions, the AI factory is specifically designed for the unique speed, scale, and complexity of modern AI workloads.

For decision-makers, vertical integration translates into more than efficiency alone. It gives organizations:

- Greater control over their own destiny
- Reduced risk and fewer points of failure
- Improved performance and reliability

## The bottom line:

Because vertical integration is so tied to tangible business benefits — control, speed, risk mitigation — it's become a compelling, highly valued approach. For the leaders reshaping AI, controlling more of the stack is the surest way to build with confidence, adapt quickly, and unlock the full promise of AI.

# 4

# Expert support fills knowledge gaps

As organizations accelerate their AI ambitions, many are coming up against a lack of internal technological expertise. In our research, decision-makers consistently cited complexity in AI model development and deployment as their greatest roadblock, closely followed by a lack of skilled AI talent on their teams. One respondent explains: “The biggest barrier is a lack of skilled AI talent. This is causing a delay in our AI adoption, impacting our ability to innovate and keep pace with competitors.”

Most organizations haven’t yet built out the deep in-house expertise required to develop, deploy, and maintain AI at scale. The challenge is compounded by a crowded ecosystem of tools and services, where even experienced technical teams can struggle to identify best practices or troubleshoot complex issues.

This is where the role of the cloud provider is being fundamentally redefined. Expert-level, proactive, and consultative support is a critical differentiator and, in some cases, the deciding factor when selecting or retaining a provider. Many organizations in our study have dropped vendors who delivered on technology but failed to deliver on partnership and support.

# Decision-Makers' Barriers for AI-Driven Initiatives

The numbers make the stakes clear:

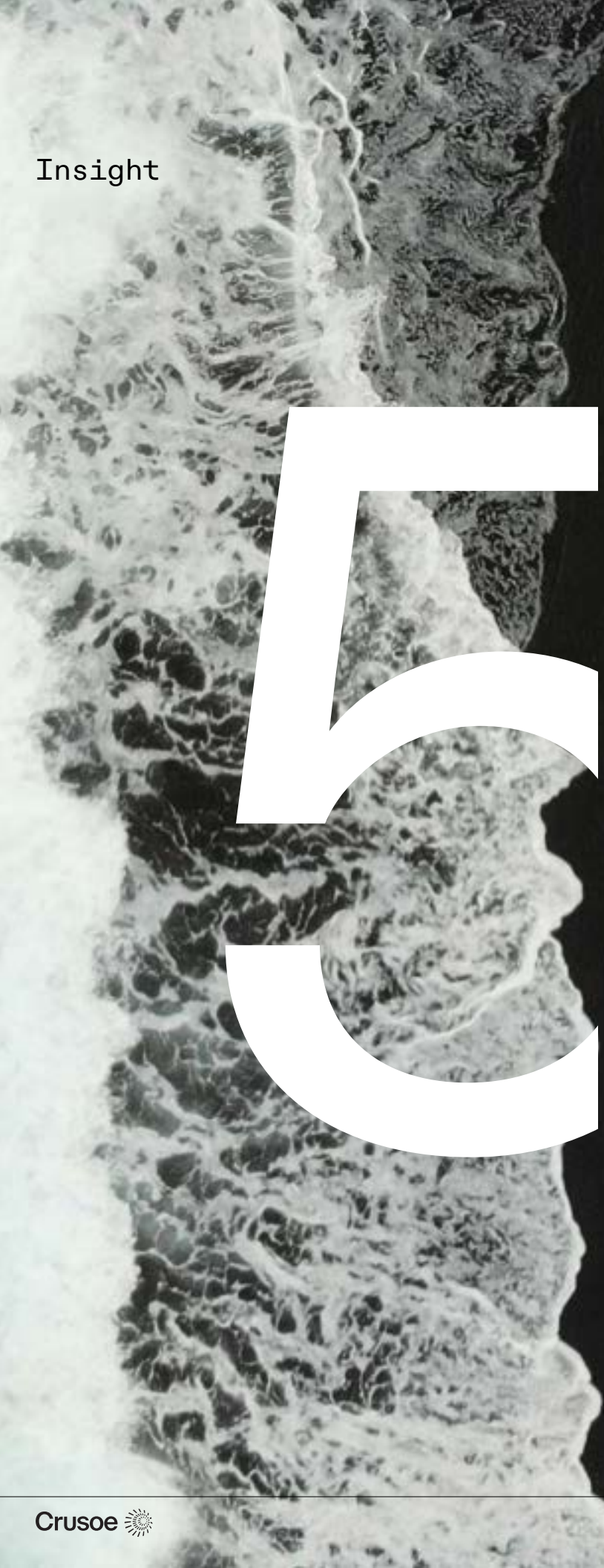
- "Complexity of AI model development and deployment" ranked as the top challenge, with an average relevance score of 5.40 out of 10.
- "Lack of internal AI expertise" closely followed at 5.16.
- In open-ended responses, decision-makers repeatedly emphasized the need for "knowledgeable guidance," "proactive insight," and "real-time troubleshooting" — often placing as much value on expert support as on technology features.

One leader states: "We've dropped cloud providers where the tech was better, but their team wasn't good enough. They couldn't keep up with us. We expect thought leadership and knowledgeable support."

Ultimately, leaders need more than a help desk. They want a true partner who can proactively identify risks and opportunities, guiding teams through unfamiliar territory. They value deep, domain-specific expertise and real-world best practices. And they need 24/7, high-touch support, not just ticket-based troubleshooting.


## The bottom line:

Expert support is critical to a successful AI implementation. The organizations that succeed will be those that surround themselves with the right mix of technology and partnership, bridging knowledge gaps and empowering teams to move from pilot to production with confidence.



# 5 Adaptability and ease of use are requirements

As teams deploy more advanced AI models and workflows, complexity has become a stubborn roadblock. Across our research, adaptability and ease of use emerged as high-impact criteria, especially for companies juggling legacy systems, new cloud-native workloads, and hybrid operating environments.

Ease of use is the most frequently cited experiential gap in hyperscaler solutions. In open-ended survey responses, decision-makers flagged steep learning curves, convoluted user interfaces, and the need for specialized training as major friction points. “We’ve definitely felt a gap when it comes to ease of use. As much as we love the scalability and features, we’ve struggled with the complexity of certain services,” one respondent explained.

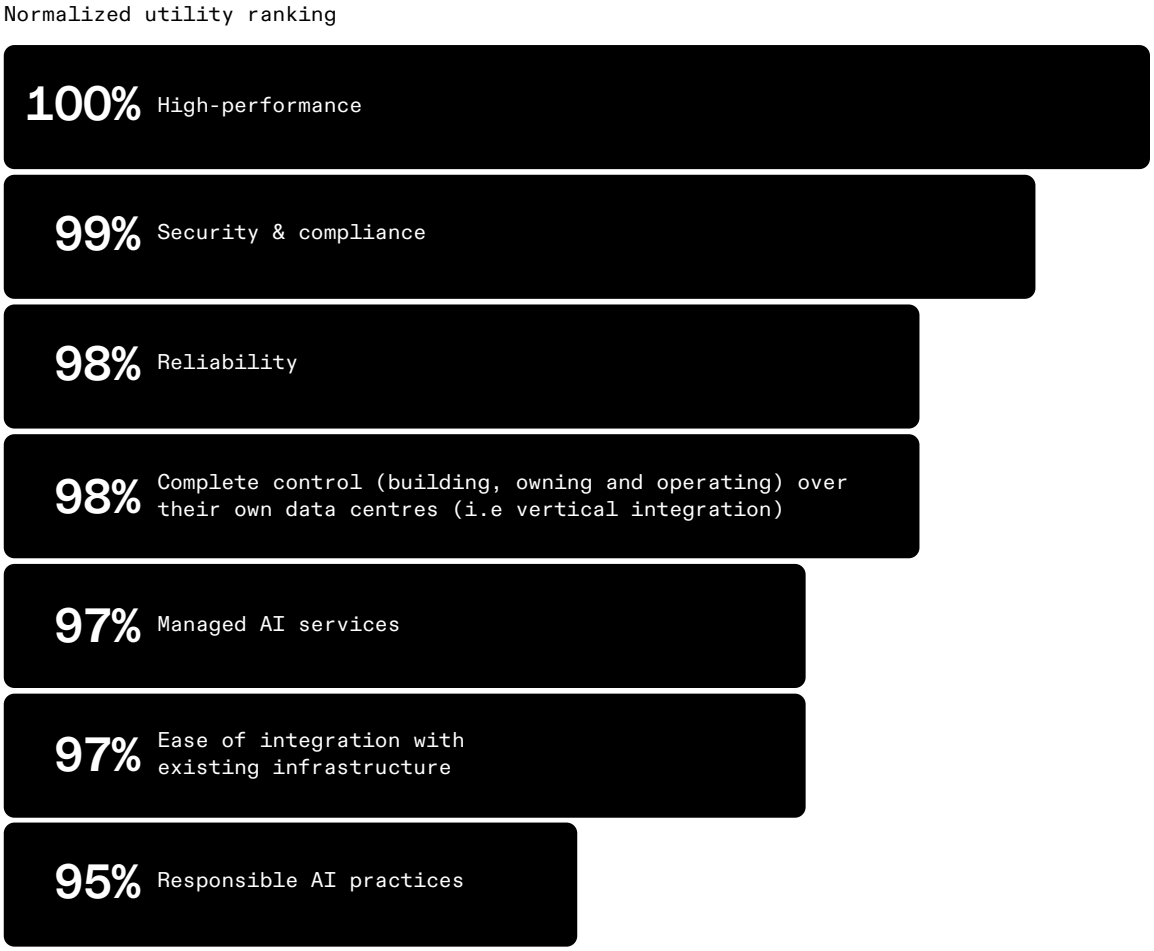
Seamless integration with existing infrastructure is another top priority — particularly for enterprises. Enterprise decision-makers ranked integration as the fifth most important attribute when choosing a provider (DNBs ranked it eighth), underscoring the challenges of connecting new AI platforms with established, mission-critical systems. Compatibility issues, delayed rollouts, and the need for custom workarounds can all stall progress and inflate costs. One leader put it simply: “The biggest gap I have with AI cloud providers is the gap between expectations and reality in terms of ease of use and support for specific AI needs.”

# Importance of cloud provider attributes for decision makers

**Key findings:**

- "Ease of integration with existing infrastructure" is rated as important by 97% of all decision-makers. For decision-makers with lower AI confidence or less internal expertise, the demand for adaptability is even greater. These leaders want more industry-specific support and tools to bridge gaps in technical knowledge and reduce their teams' reliance on outside consultants.
- Adaptable solutions that can flex across industries and use cases are also highly valued, with 40% of respondents saying they require more specialized support or customization to meet the unique demands of their sector.

As AI complexity grows, leaders are searching for solutions that take friction out of the equation. They need tools that are intuitive to use, integrate seamlessly with existing systems, and adapt to the unique demands of their industry and teams.




## The bottom line:

Adaptability and ease of use are no longer "nice to have" features. They are essential for reducing complexity, accelerating time-to-value, and making AI accessible to the teams that will drive the next wave of innovation.

# 6 Responsible AI and sustainability are rising priorities

As AI's reach expands, so do expectations for responsible use and mitigating environmental impact. While cost, performance, and reliability still top the list for most organizations, our research shows that responsible AI and sustainability are rising rapidly as decision criteria — especially for digital-native businesses and ESG-driven organizations.

For DNBs, "responsible AI practices" ranked as the fifth most important attribute when evaluating cloud providers (out of 15 total), while for enterprises, it ranked eighth. Many respondents predict that as their AI initiatives mature — and as regulatory and ESG frameworks evolve — these considerations will become essential, not optional.

> I will say, responsible AI practices are of emerging importance to us. We want to make sure there aren't any destructive or negative consequences to the work we are putting out [to the world].

Decision-Maker
CEO & Co-founder
Real Estate Franchise

---

A key aspect of this shift is transparency around energy sourcing. Increasingly, builders expect providers to clearly demonstrate how AI workloads are powered, particularly with regard to renewable energy. Dissatisfaction is growing with vague or unsupported sustainability claims. For many companies and ESG-focused organizations, energy sourcing is now viewed as an extension of responsible technology, and, in some cases, a tiebreaker. As one leader put it: "To guarantee responsible and ethical development and deployment of AI [solutions] at [our organization], we will always consider providers' practices."

The call for clarity, specificity, and proactive commitment is unmistakable. Responsible AI practices are increasingly a tiebreaker — and, for some organizations, a prerequisite for consideration at all.

## The bottom line:

Responsible AI and sustainability may not be the first concern for every organization today. But as expectations rise and the market matures, leaders recognize that these values are fast becoming a defining standard that will shape both competitive positioning and societal trust in the era of AI.

# AI will provide the edge, and infrastructure will decide who wins

As organizations move from pilot projects to business-wide AI integration, the question becomes how quickly and how sustainably teams can realize a competitive edge.

Seizing the full potential of AI requires a new kind of infrastructure partner — one built for the challenges and opportunities of this era.

While legacy hyperscalers are working to adapt existing infrastructure to meet AI's demands and many neoclouds are struggling to deliver at scale, a new approach is needed. Every layer of the stack must be designed specifically for AI workloads and optimized for the unique demands of generative models, deep learning applications, and intelligent tools.

## Control your future with vertical integration

Decision-makers in our study voiced a strong desire for greater command over their technology environments. Vertical integration — combining energy sourcing, data center construction/operations, and high-performance cloud delivery into a unified system — has emerged as a powerful model to address this need.

**In practice, vertical integration enables:**

- Direct oversight of performance and security: no more waiting on external vendors to resolve issues
- Faster response to opportunity and change: your infrastructure can adapt as quickly as your strategy
- Reduced risk: fewer points of failure and greater supply chain resilience

## More than raw compute

While access to compute power is critical, it is not, by itself, a differentiator. Our findings show that speed to market — the ability to rapidly deploy, iterate, and scale — is the real advantage in today's environment. Modular, quickly deployable data center solutions (such as Crusoe Spark™) exemplify this shift: enabling organizations to expand capacity in weeks rather than months, giving teams the agility to seize AI opportunities as they emerge.

## Simplicity meets expertise to unlock possibilities

For many organizations, the complexity of deploying and scaling AI has created exclusion rather than empowerment. Teams describe feeling "locked out" by technical hurdles or unsupported by generalized solutions. Our research underscores that reducing complexity and coupling technical depth with high-touch guidance are key to enabling broader participation and faster progress with AI.

## Commitment to responsible AI

Leaders increasingly recognize that true progress in AI must be measured not only by technical achievement, but also by the broader consequences for people and the planet. Decision-makers are asking tough questions about how their AI initiatives affect privacy, equity, and environmental stewardship. The organizations setting the pace increasingly expect partners to:

- Demonstrate clear ethical guidelines and guardrails for AI development and deployment, reducing risks of bias, misuse, and unintended harm.
- Invest in sustainable operations and renewable energy to limit environmental footprint and align technology growth with climate goals.
- Provide transparency and accountability in how data is sourced, how models are trained, and how decisions are made at every level.

# Building the future faster: Partnering for transformative growth

AI is now the defining force in business transformation. But building the future requires a new foundation designed specifically for AI-powered innovation.

## The vision

Crusoe is driven by a bold mission: to accelerate the abundance of energy and intelligence for a world hungry for progress. We see AI not just as a tool, but as a catalyst for new opportunity, enabling organizations to turn ambitious ideas into reality.  And we believe that with the right infrastructure, every organization can make that transformation faster and easier, with less risk.

## The Crusoe difference

What sets Crusoe apart is our AI factory model — a vertically integrated, energy-first, purpose-built approach to AI infrastructure:

- Climate-aligned energy: Crusoe harnesses climate-aligned energy, reducing environmental impact while lowering long-term costs.
- Next-generation data centers: We design, build, and operate our own high-performance data centers, setting records for speed to market and reliability.
- AI-optimized cloud: Our platform is built from the ground up for demanding AI workloads, offering unmatched control, transparency, and performance.
- Modular scalability: With solutions like Crusoe Spark™, organizations can rapidly deploy and scale high-performance AI capacity — wherever and whenever it's needed.
- Consultative support: Crusoe delivers 24/7 global support with industry-leading satisfaction, ensuring every customer has access to expert guidance and proactive partnership. Our average reply time is just under six minutes, with average resolution in 23.6 hours. We've had a consistent 100% customer satisfaction rating since June 2024.

Crusoe

# Build the future faster

Contact us →

### The path ahead

For companies ready to build ambitiously with AI, Crusoe is the essential partner for the journey ahead. Our seamless platform — integrating climate-aligned energy, advanced engineering, and a high=performance cloud purpose-built for AI — empowers your teams to conquer technical and operational barriers. The winners of tomorrow are building differently, starting today.

Book a call to learn how you can build the future faster, smarter, and more sustainably with Crusoe.



Privacy Policy

Crusoe is an NVIDIA Preferred Partner