State of AI
Infrastructure Report

2025

flexential.com

Table of contents

INTRODUCTION

Companies see the use of AI as essential to staying competitive, but have they built the right systems and processes to manage the demands that come with it?

AI is no longer a tentative or temporary experiment for most organizations. It has become a fundamental part of business operations, shaping strategy, decision-making, and innovation. But while AI's presence is expanding, so are its complexities.

For our second annual State of AI Infrastructure Report, **we surveyed over 350 IT leaders** at companies with more than $100 million in annual revenue—including 100 respondents from organizations exceeding $2 billion—to understand how businesses are implementing AI and how it's affecting their IT infrastructure.

What's clear is that companies are increasingly investing in AI technologies and expect to see measurable returns in short order. **Nine in 10 companies (90%) are deploying or planning to deploy generative AI,** and **more than half of respondents are using it for predictive analytics, cybersecurity, autonomous systems, computer vision, or natural language processing (NLP)-based applications [Fig. 1].**
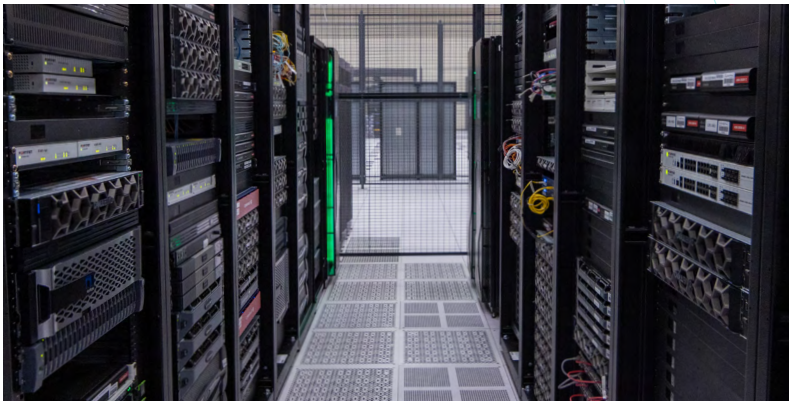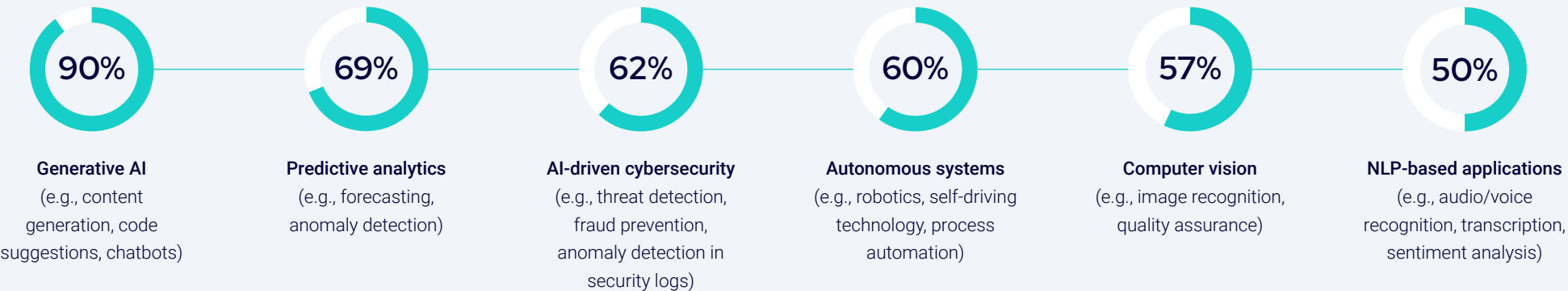
FIG. 1

What types of AI/ML use cases are you deploying or planning to deploy?

90%
Generative AI
(e.g., content generation, code suggestions, chatbots)

69%
Predictive analytics
(e.g., forecasting, anomaly detection)

62%
AI-driven cybersecurity
(e.g., threat detection, fraud prevention, anomaly detection in security logs)

60%
Autonomous systems
(e.g., robotics, self-driving technology, process automation)

57%
Computer vision
(e.g., image recognition, quality assurance)

50%
NLP-based applications
(e.g., audio/voice recognition, transcription, sentiment analysis)

Only 5% of organizations describe their AI adoption as nascent (down from 10% a year ago), and **optimism about AI's use remains high**—three-quarters (75%) of IT leaders express excitement about AI's role in their organizations [Fig. 2 and 3]. **However, the number of respondents feeling overwhelmed by AI's implementation has more than doubled since last year, from 12% to 29% [Fig. 3].**

In addition, planning ahead has become essential. Most organizations (62%) are

FIG. 2

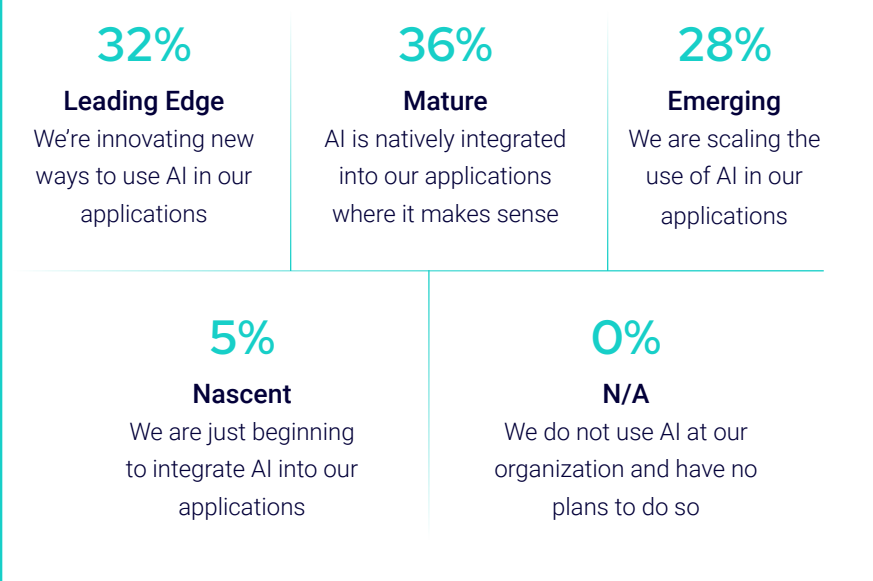*Which of the following best describes the state of artificial intelligence (AI) at your organization?*

32%	36%	28%
Leading Edge	**Mature**	**Emerging**
We're innovating new ways to use AI in our applications	AI is natively integrated into our applications where it makes sense	We are scaling the use of AI in our applications

5%	0%
Nascent	**N/A**
We are just beginning to integrate AI into our applications	We do not use AI at our organization and have no plans to do so

FIG. 3

Which of the following best describes your attitude toward the implementation of AI applications and initiatives in your organization?

Excited

75%

Proud

55%

Inspired

47%

Overwhelmed

29%

Nervous

13%

Apathetic

12%

Uncertain

10%
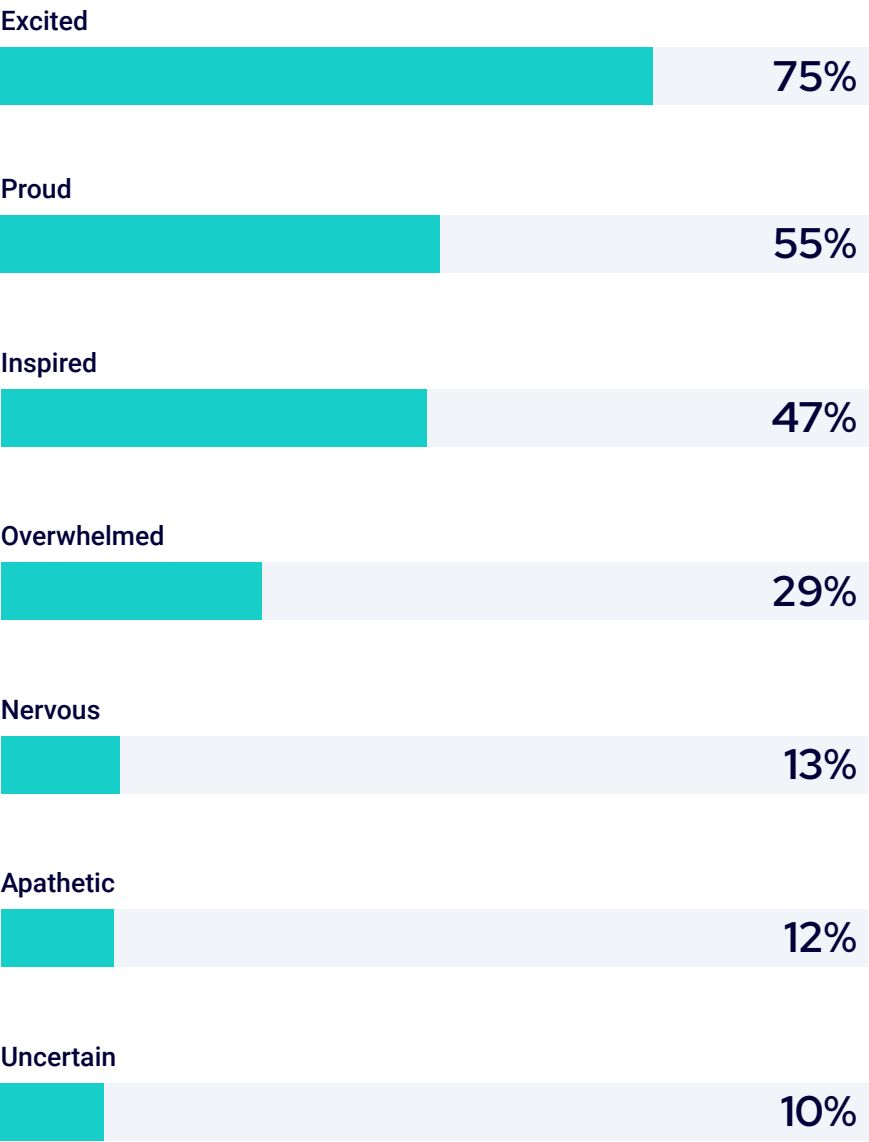
In addition, planning ahead has become essential. Most organizations (62%) are **mapping out their IT infrastructure and data center capacity needs one to three years in advance**, with another 17% looking three to five years ahead. Despite the tight timelines and rising demands, 94% of respondents **expressed confidence in their planning process**. Even among those with less than a one-year horizon, a surprising 70% said **they feel well prepared to meet future infrastructure requirements** though vacancy rates are at a record-low.

Clearly, while companies see AI's value, they are also **grappling with its demands**—that show no sign of slowing—and will need to plan accordingly.

Five biggest takeaways

While organizations are bullish on AI and increasing their supporting investments, they also:

1 Expect rapid financial returns on AI spending

2 Struggle with infrastructure constraints that hinder expansion

3 Face growing skills shortages in AI implementation

4 Rely on inadequate data center planning cycles

5 Encounter network performance and security issues that limit scalability

Eight key findings

53% → 71%

Confidence in organizations' **ability to execute their AI roadmaps** has grown significantly, **rising from 53% to 71%** in one year.

53% → 81%

of respondents said the **C-suite is the driving force behind their organization's decision** to adopt AI-driven applications. At an **increase of 28 percentage points**, there is considerably more buy-in than a year ago.

70%
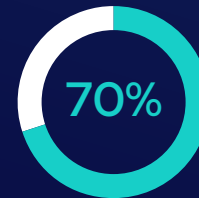
of respondents reported devoting at least **10% of their organization's total IT budget to AI initiatives**, including software, hardware, and networking.

44%
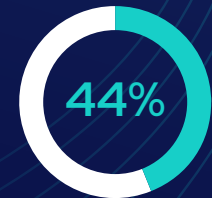
of respondents said **IT infrastructure constraints are the greatest barrier** to expanding their organization's AI initiatives.

62%
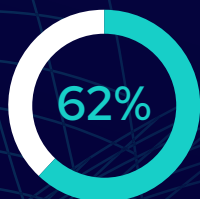
of respondents reported **planning their IT infrastructure and data center capacity needs** one to three years in advance in response to **increased demand** and **limited immediate availability**.

33%
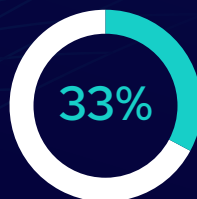
of respondents said their **organization's AI governance policies don't cover security protocols for AI systems and data**, and nearly half (48%) reported gaps in policies addressing bias detection and mitigation in AI models.

86%
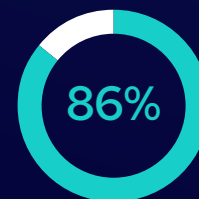
of respondents are **worried about acquiring or developing the specialized talent** needed to meet AI goals.

53% → 61%

61% of respondents have **encountered skills or staffing gaps** in the management of **specialized computing infrastructure**, up from 53% a year ago.

SECTION 1: THE AI ROADMAP

C-suite leadership drives investment as AI becomes a core business strategy

Just last year, AI adoption within organizations was still largely exploratory, with leadership weighing its potential against implementation challenges.

Now, **AI has moved beyond experimentation** as senior executives are integrating it into core business strategies at an accelerated pace. **The C-suite has clearly become the driving force behind AI initiatives** in the past year, with 81% of respondents citing the highest level of corporate leadership as the reason why this technology is being implemented throughout the organization [Fig. 4]. **It's a significant increase from last year**, when only 53% of respondents credited the C-suite with leading AI adoption.

Still, senior executives **hesitate to commit to bold strategies without a proven return on investment**, **industry validation**, and **risk mitigation**. High-profile AI implementation failures, like infrastructure bottlenecks and latency issues, have made them wary, as have rogue behaviors and AI hallucinations. Meanwhile, issues like model bias, missteps in decision-making, and regulatory risks **demand strong oversight**. However, as AI's business value becomes clearer, leadership is taking a more active role in funding its success.

81% of respondents credited the **C-suite** with leading AI adoption
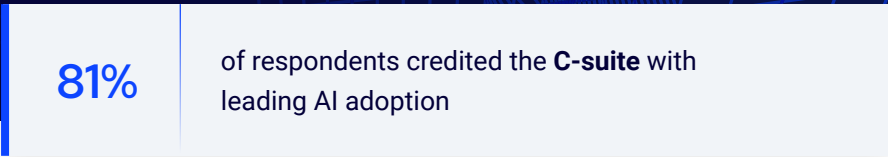
FIG. 4

Which of the following groups or individuals are the driving force behind your organization's decision to adopt AI-driven applications?

C-Suite	Board	Myself (IT leaders)	Competitors (i.e. we need to keep up with our industry)	Our customers	Our employees	The general public
81%	46%	34%	33%	26%	19%	13%
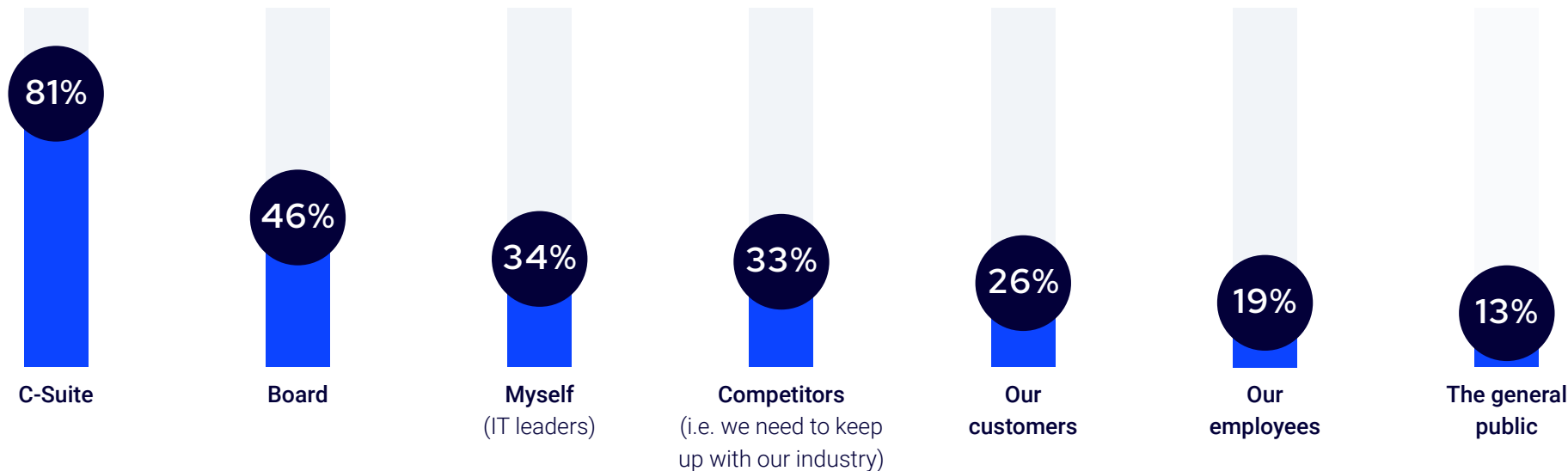
AI spending reflects growing importance

AI investment accounts for an increasing portion of IT budgets. Seven in 10 respondents said their organization allocates at least 10% of its IT spending to AI-driven software, hardware, and networking. A quarter are dedicating more than 20% of their budgets to AI initiatives, underscoring the technology's importance to long-term growth [Fig. 5].

Billions are already being invested in **next-generation data centers**, **cloud infrastructure**, and **AI-specific hardware**. With so much capital at stake, investors and boards expect quick financial payoffs to justify these expenses. More than half (51%) of organizations anticipate measurable financial benefits within the next year, and another 23% expect it within one to three years [Fig. 6].

These expectations align with how companies define AI's success: 58% measure the ROI of AI initiatives in terms of revenue growth and market share, while 50% prioritize cost reduction and operational efficiency. And when companies using generative AI are reporting an average ROI of $3.70 for every dollar spent,[1] it's clear why C-suite leaders are increasingly committing to AI [Fig. 7].

FIG. 5

Approximately what percentage of your total IT budget is allocated to AI initiatives (software, hardware, and networking)?

More than 20%

25%

10-20%

45%

5-10%

23%

Less than 5%

7%

FIG. 6

When does your organization expect to generate measurable financial benefits from its AI investments?

21% We are already seeing financial benefits

51% Within the next year

23% In one to three years

5% In three to five years

0% More than five years from now

0% We do not expect AI to generate direct financial returns
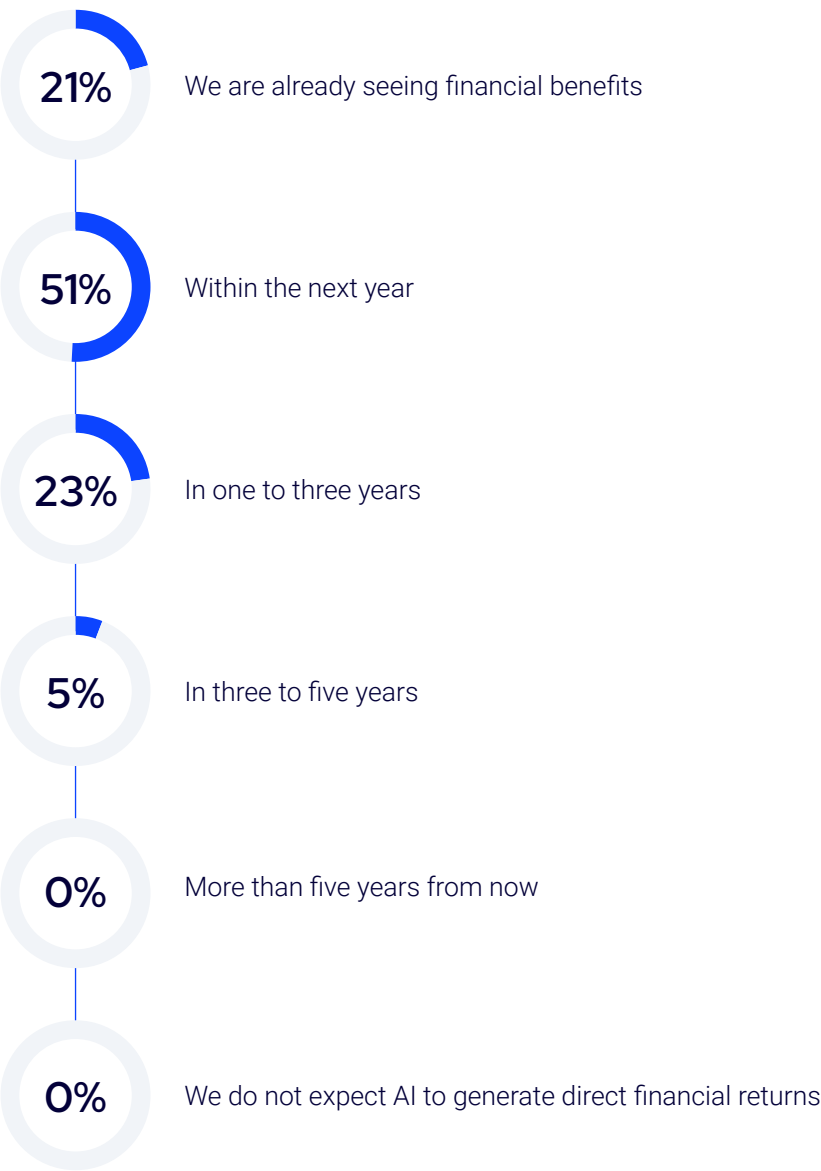
FIG. 7

How do you measure the return on investment (ROI) of your AI initiatives?

58%	**50%**	**47%**	**40%**	**35%**	**26%**	**20%**	**0%**
Revenue growth or increased market share	Cost reduction/ operational efficiency	Customer satisfaction metrics (Net Promoter Score, churn rate)	Time to market or process improvement	Intangible benefits (brand reputation, customer loyalty)	Monetization (creating new revenue streams from AI-generated insights)	Employee productivity and augmentation improvement	We don't formally measure ROI for AI initiatives

Competitive pressure, market forces accelerate AI adoption

AI adoption is not happening in isolation. Companies are **under increasing pressure** from **competitors**, **investors**, and **market dynamics** that reward efficiency and profitability. Plus, tech giants like Microsoft, Google, and OpenAI are investing billions into AI-powered products and services, setting high expectations for adoption. Businesses that **hesitate risk falling behind** others that are **already capitalizing on AI-driven gains**.

Hype and success stories fuel adoption. A significant 92% of organizations acknowledge that media coverage, industry trends, and technological breakthroughs influence executive buy-in for AI projects [Fig. 8]. High-profile success stories—such as Walmart's use of large language models (LLMs) to optimize product data—**demonstrate AI's impact on efficiency and customer experience**.[2] When industry leaders showcase AI-driven improvements, it **sets new benchmarks for success** and **compels others to invest**. Executives feel pressured to keep pace to avoid missing out on market opportunities.

At the same time, financial markets **reward companies that can demonstrate early wins**. Investors expect AI investments to translate into revenue gains and cost reductions. Companies like DeepSeek, which developed a top-performing reasoning model[3] with significantly less funding (though potentially by cutting corners),[4] **highlight AI's ability to drive efficiency** at a lower cost.
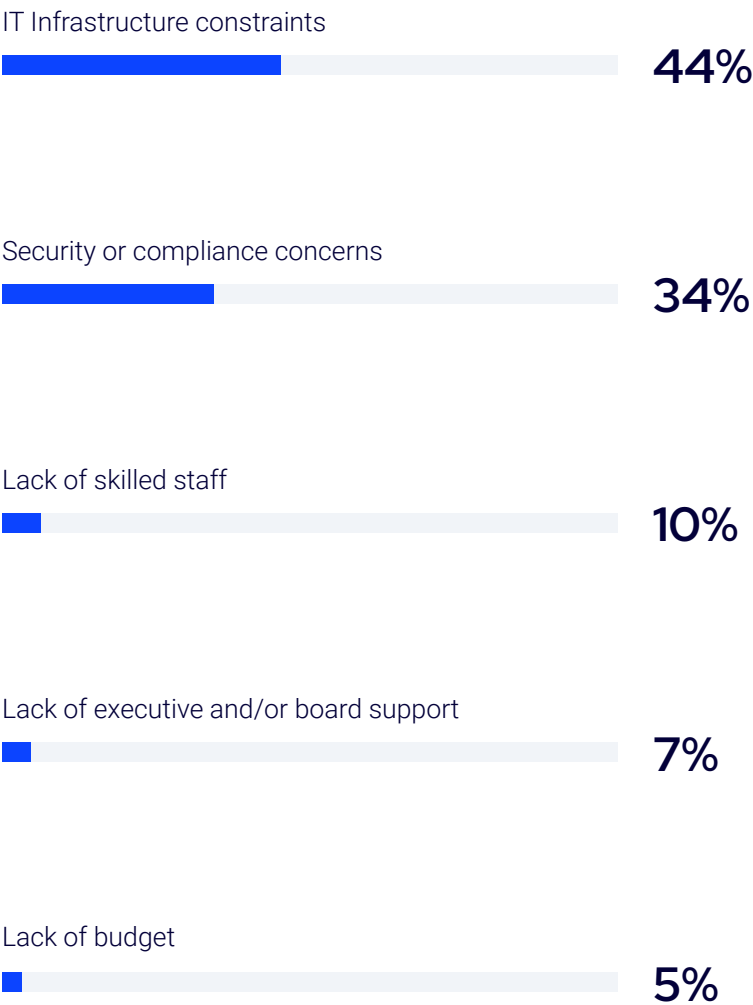
Organizations that can deliver measurable results through automation, data-driven decision-making, or enhanced customer interactions are better positioned to secure continued funding and executive support.

FIG. 8

To what extent do external forces, including media coverage, industry trends, and technology breakthroughs, influence internal stakeholder support or executive buy-in for AI projects?

Significantly — External forces are a primary driver of support and decision-making.

30%

Moderately — These factors are influential, but internal assessments and strategies are more pivotal.

62%

Minimally — While aware of external influences, decisions are predominantly based on internal evaluations.

8%

Not at all — We base our decisions strictly on internal criteria and strategic goals, independent of external forces.

1%

FIG. 9

Which of the following represents the greatest barrier to expanding your organization's AI initiatives?

IT Infrastructure constraints

44%

Security or compliance concerns

34%

Lack of skilled staff

10%

Lack of executive and/or board support

7%

Lack of budget

5%

IT infrastructure constraints pose the biggest challenge

While enthusiasm for AI adoption is high, implementation is not without challenges. IT infrastructure constraints are the **single greatest barrier to scaling AI initiatives**, cited by 44% of organizations [Fig. 9]. AI workloads demand significant computational power, and many legacy systems are not built to support high-density processing.

Data centers, cloud capacity, and networking **must all evolve** to meet **AI's increasing requirements**. Organizations recognize these limitations but view them as solvable—not by fixing everything in-house, but by **turning to partners that offer access to scalable power**, **advanced cooling**, **high-density capacity**, and **purpose-built facilities**. Strategic investments in next-generation data centers, hybrid cloud environments, and colocation solutions are seen as necessary steps to unlock AI's full potential.

Security and compliance concerns also remain a challenge. As AI adoption grows, so do concerns over **data privacy**, **regulatory alignment**, and **model governance**. Executives need clear frameworks to ensure AI systems operate within ethical and legal boundaries. Without these assurances, scaling AI initiatives will be difficult.

AI's momentum continues to build

With strong executive buy-in, rising investment levels, and clear business outcomes, **AI has become a competitive necessity**. The rapid pace of AI innovation and the success stories emerging across industries, including **financial services** (AI transcription saves financial advisers 10 to 15 hours each week),[5] **mining companies** (automating administrative tasks saves 2,200 hours a month),[6] **public relations firms** (productivity is up 10.2%),[7] and **health care** (saving 11,000 nursing hours and nearly $800,000),[8] are reinforcing confidence in its ability to drive tangible results.

AI investments are also resonating with consumers, who expect **personalized offers** and **round-the-clock support**. To meet these demands, businesses are leveraging AI to analyze vast data sets and deliver actionable insights at a scale beyond human capacity.

Organizations that have integrated AI into their operations are **seeing measurable gains in efficiency and profitability**. But as AI adoption accelerates, executives' focus is shifting from proving AI's value to ensuring their company doesn't fall behind.

Key takeaways

1

AI adoption has shifted from exploration to execution, with the **C-suite now spearheading investment decisions** and prioritizing measurable returns. Accordingly, organizations are dedicating larger portions of their IT budgets to AI, with 51% expecting financial benefits within a year and the majority evaluating success through revenue growth, cost savings, and efficiency gains.

2

Executives view **infrastructure constraints as the primary barrier to AI expansion**, but they're viewed more as challenges to overcome rather than considerable roadblocks. Legacy systems and scalability issues are slowing deployment, prompting investments in next-generation data centers, hybrid cloud environments, and AI-ready infrastructure to support growing computational demands.
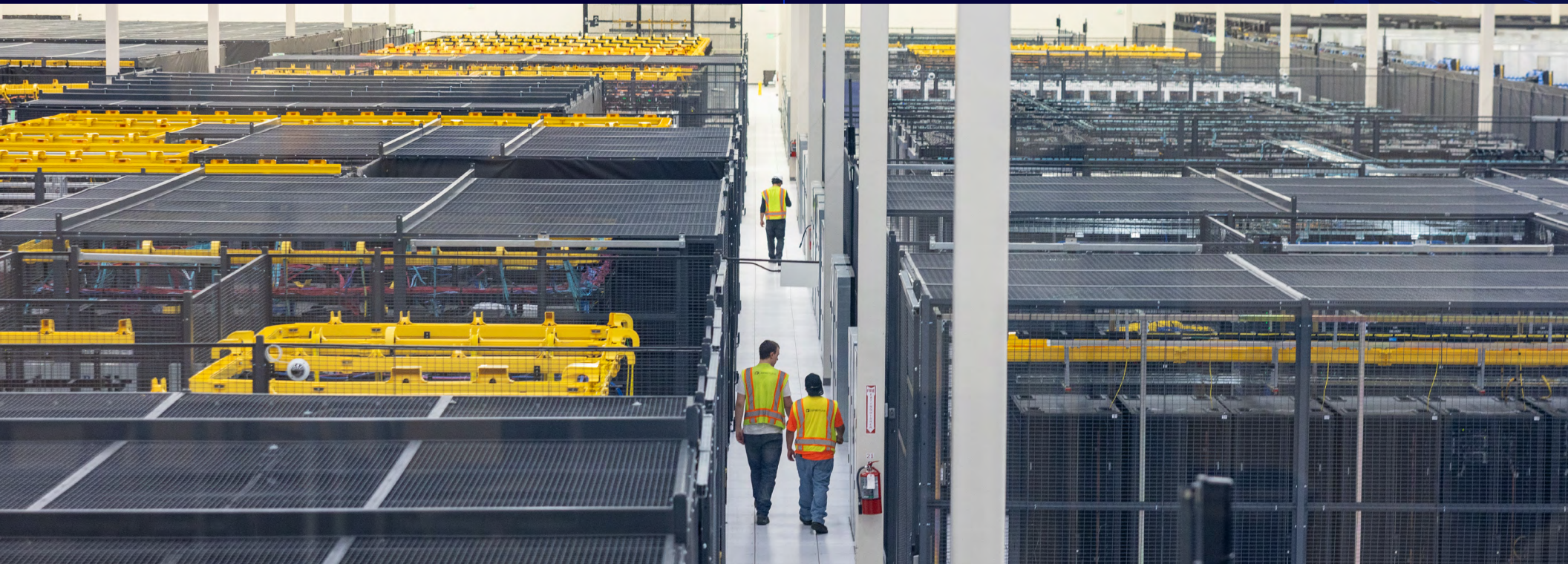
SECTION 2: LEADERSHIP INFLUENCE

Executives grow more confident in AI roadmaps as investments pay off

A year ago, just over half (53%) of executives were confident their company could execute its AI roadmap.

Today, that number has increased from 34% to 71%, signaling a shift from cautious exploration to full-scale deployment [Fig. 10].

Several factors underpin this increase in confidence. Investment in AI infrastructure has accelerated, with enterprises spending $246 billion on AI solutions last year.[9] This backing means **more data centers are optimized for AI workloads**, offering enterprises more scalable computing power than ever before.

At the same time, AI tools—particularly generative AI—have demonstrated immediate value. Business adoption of generative AI jumped from 55% in 2023 to 75% last year. That has fueled employee use: The percentage of professionals using AI tools at least weekly jumped from 20% in July 2023[10] to 30% in late 2024.[11]
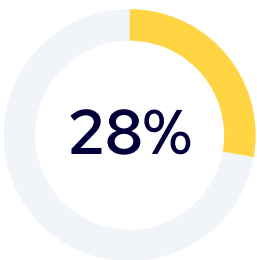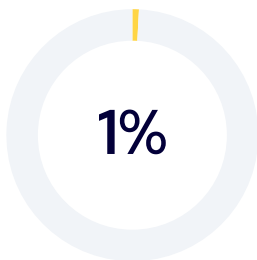
FIG. 10

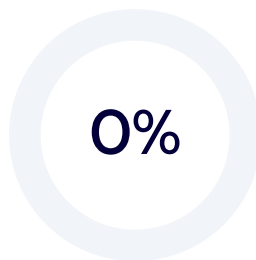*Which of the following best describes your confidence level in your organization's ability to execute its AI roadmap?*
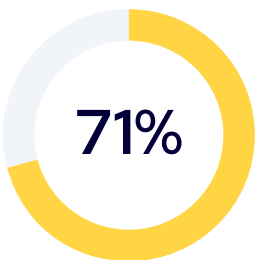
71%	28%	1%	0%
I'm extremely confident	I'm somewhat confident	I'm not very confident	I'm not at all confident

The competitive stakes of AI execution

While confidence in AI has grown, so have the **consequences of falling behind**.

Among executives we surveyed, 28% cited losing market share as the most significant risk if they fail to meet their AI goals. Another 26% pointed to delays in bringing AI-driven products to market. Revenue concerns are also prominent, with 17% identifying financial underperformance as a direct consequence of failing to execute on their AI strategy [Fig. 11].

The urgency to deploy AI effectively is also evident in how companies measure success. As noted, 51% of organizations expect to see measurable financial benefits from AI within the next year, and 21% report that they already are.

These expectations indicate that **AI is not being treated as an experimental investment**. It is expected to deliver results quickly. Workforce considerations further amplify the pressure to act: 10% percent of respondents expressed concern about losing employees to competitors with more advanced AI strategies, underscoring the role AI plays in talent attraction and retention.

FIG. 11

What is the most impactful consequence if your organization does not achieve the goals laid out in its AI roadmap?

We will lose market share to competitors

28%

We will have to delay time to market for new products or services

26%

We will not meet our revenue targets

17%

We risk not being compliant with regulations

12%

We will lose employees to companies with more advanced AI strategies

10%

We won't be able to meet customer demands

7%

FIG. 12

Which areas are covered by your organization's AI governance policies?

Security protocols for AI systems and data

67%

Data privacy and user consent management

58%

Transparency and explainability of AI decision-making

58%

Bias detection and mitigation in AI models

52%

Vendor risk management for third-party AI solutions

41%

Continuous monitoring and updating of AI systems and policies

29%

We do not have specific governance measures in place

0%
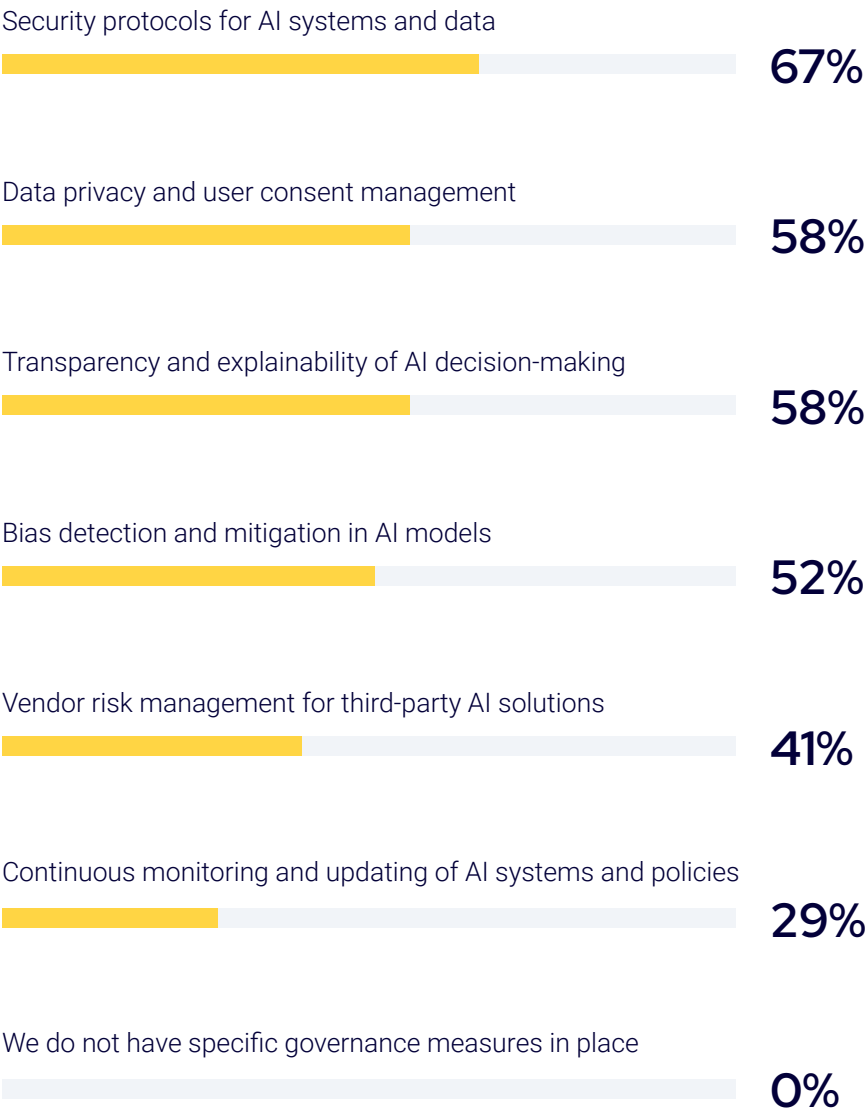
Confidence, governance go hand in hand

The increase in confidence is a reflection of progress—and necessity.

AI is becoming a defining factor in market leadership. The focus now is shifting to whether organizations can quickly implement it on a larger scale.

Stronger AI governance frameworks reinforce this confidence. Concerns about **security**, **compliance**, and **regulatory challenges** have been perceived as barriers to AI adoption, but improved risk mitigation strategies have eased these fears.

Two-thirds (67%) of companies have established security protocols for AI systems and data, while 58% have frameworks for data privacy, user consent management, and transparency in AI decision-making [Fig. 12]. This highlights how **companies are taking a risk-based and measured approach** as necessary in their AI adoption.

Bias detection and mitigation are also priorities, with 52% actively addressing these concerns in their AI models. With clearer legal frameworks and more robust security protocols in place, **businesses feel more assured that they can integrate AI** while **remaining compliant and secure**.

Better-defined regulatory frameworks and improved data protection strategies have **reduced many of the risks** that previously held companies back from fully adopting AI. With these foundations in place, the next challenge will be scaling AI effectively while maintaining security and compliance.

Key takeaways

1

Confidence in organizations' **ability to execute their AI roadmaps** has grown significantly, rising from 53% to 71% in one year, as businesses move from experimentation to measurable results. This shift is driven by increased investment in AI infrastructure, stronger governance frameworks, and early financial returns from AI deployments.

2

The risks of **failing to meet AI goals** have **become more defined**, with companies citing market share loss, delayed product timelines, and revenue shortfalls as key concerns. At the same time, governance strategies are maturing, with most organizations implementing security, transparency, and bias mitigation policies to ensure responsible AI adoption.

SECTION 3: THE AI SKILLS GAP

Workforce shortages threaten AI growth and business innovation

As AI systems grow more complex, businesses are struggling to find employees with the skills to manage them.

Only 14% of respondents believe they have the right people in place to help them meet their AI goals, and workforce limitations have only **worsened** in the past year.

We found that 61% of organizations reported skills shortages in managing specialized computing infrastructure this year, up from 53% last year, **highlighting the increasing strain** on IT teams as AI deployments grow in scale [Fig. 13].

After 39% of companies reported shortages in data science and data engineering roles last year, 53% said they had encountered those issues this year.

These are **coveted roles** that are **not being filled**. In the long run, this lack of skilled workers can stall AI projects—if they even get out of the sandbox.
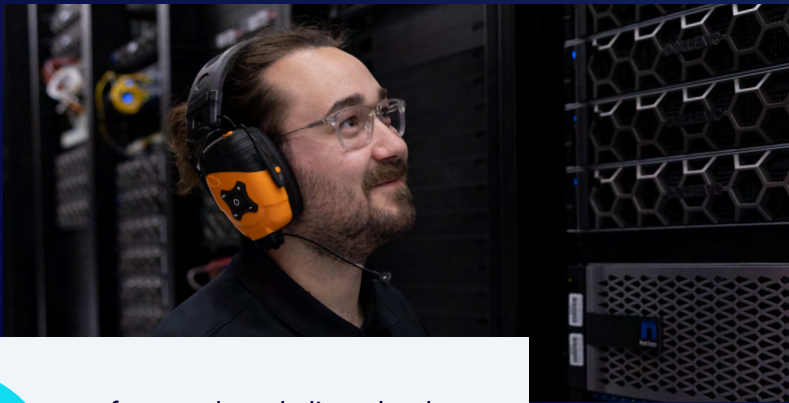
14%

of respondents believe they have the **right people in place to help them meet their AI goals,** and workforce limitations have only worsened in the past year.

FIG. 13

In the past year, has your organization encountered skills or staffing gaps in any of the following areas related to AI?

61%	53%	47%	47%	5%
Management of specialized computing infrastructure (e.g., high-density computing)	**Data science or data engineering**	**Management of advanced networking technologies** (e.g., SDN, NFV)	**Cybersecurity**	**N/A** — Our team has not encountered skills or staffing gaps in the past year

The competitive market for AI talent

The demand for AI talent continues to grow, creating a **fiercely competitive hiring environment**, particularly in specialized areas such as **deep learning** and **natural language processing**. Job postings that reference AI have risen 21% annually since 2019,[12] yet the supply of skilled workers has not kept pace.

Nearly one in four U.S. tech jobs posted early this year indicate a **need for employees with AI skills**,[13] but competition for AI expertise extends beyond traditional tech firms.

Companies are pulling out all the stops to win over the limited candidates. **Compensation has surged** 11% per year,[14] with companies offering substantial salaries, bonuses, and stock options to secure top talent. **Remote work is another lure employers can use** to expand their talent pool and attract talented individuals from around the world.

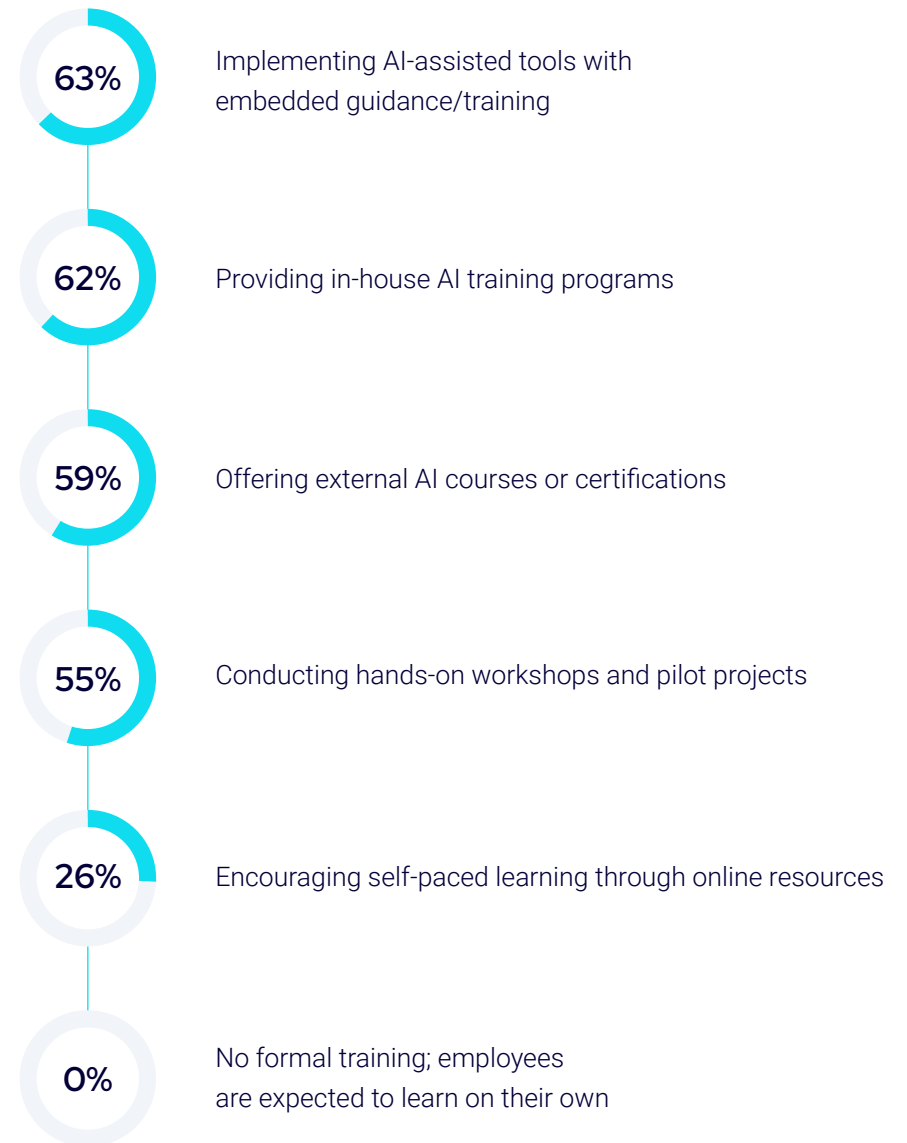## Addressing the skills gap with targeted training

Despite these challenges, organizations believe they can close the talent gap internally. Businesses are training employees by **investing in AI-assisted tools**, **internal upskilling programs**, and **external certification courses**. Notably, we found that 63% of organizations now implement AI tools with training built in to help employees develop skills in a practical context, while 62% offer structured in-house AI training programs [Fig. 14].

However, the need for retraining extends beyond technical skills. The mismatch between AI initiatives and workforce capabilities is **evident**, with 36% of companies reporting that they struggle to align staff with evolving AI demands. The pace of innovation in AI means that even **experienced professionals must continuously update their skills**.

Unfortunately, due to increased or changing workloads, **employees may not have time** for adequate upskilling. Nor does everyone receive the **same opportunity** to uplevel their AI skills: 71% of AI-skilled workers are men, and only 22% of baby boomers receive training.[15] Unless organizations broaden who they train and hire in AI, they will only **exacerbate talent shortages**.

FIG. 14

How is your organization training employees to effectively use AI in their roles?

63% Implementing AI-assisted tools with embedded guidance/training

62% Providing in-house AI training programs

59% Offering external AI courses or certifications

55% Conducting hands-on workshops and pilot projects

26% Encouraging self-paced learning through online resources

0% No formal training; employees are expected to learn on their own

The future of workforce development

When a company can't hire enough AI specialists, the few skilled employees on staff may end up feeling like they're **stretched too thin**. Teams lacking sufficient expertise may have to maintain complex AI systems with too few hands, or members may juggle a complex workload. These skills gaps then can have other **repercussions** as they may **hurt job performance** and **fuel burnout**.

Efforts to avoid this outcome have given rise to another implementation of AI. Companies are **exploring automation in talent management** by using AI tools to optimize recruitment, predict workforce needs, and personalize training programs to attempt to mitigate the gap.

These constraints highlight why businesses are **prioritizing workforce development**, even as hiring conditions shift. Without enough skilled workers, companies may struggle to realize AI's full potential.

Key takeaways

1

The AI talent shortage is one of the biggest barriers to adoption, with **organizations struggling to find employees who have the specialized skills** needed to support AI initiatives. Shortfalls in infrastructure management, data science, and cybersecurity are growing, making it harder for businesses to scale AI effectively.

2

Companies are relying on a mix of AI-assisted tools, in-house programs, and external certifications to train employees, but gaps remain, especially as **AI evolves faster than traditional education can keep up**. Without sustained investment in workforce development, businesses risk delays, inefficiencies, and underutilized AI infrastructure.

SECTION 4: CAPACITY CHALLENGES

AI expansion, limited capacity force enterprises to rethink data center infrastructure planning

Securing data center capacity has become a race against time.

Organizations are encountering a market where **demand continues to outstrip supply**, power constraints threaten expansion, and **traditional 18- to 24-month planning cycles** are **proving inadequate**. The shift is particularly urgent for enterprises relying on AI-driven workloads, where failure to secure infrastructure well in advance could leave them without the resources to scale.

Yet, despite clear market indicators, nearly all respondents (94%) expressed **confidence** in their planning process [Fig. 15]. The 17% of respondents looking

three to five years ahead are in the strongest position to secure capacity, while the 62% who plan their needs one to three years ahead face more competition and risk [Fig. 16].

Of the 16% of organizations planning less than a year ahead, 70% still expressed **strong confidence** in their ability to meet future IT infrastructure and data center capacity needs. That belief may be misplaced: Vacancy rates in primary data center markets fell to a record-low 1.9% at the end of last year.[16] And, companies seeking more than 5 MW of capacity now face **wait times of up to 24 months** because of low vacancy rates and more than 70% of new colocation builds being preleased.[17]

FIG. 15

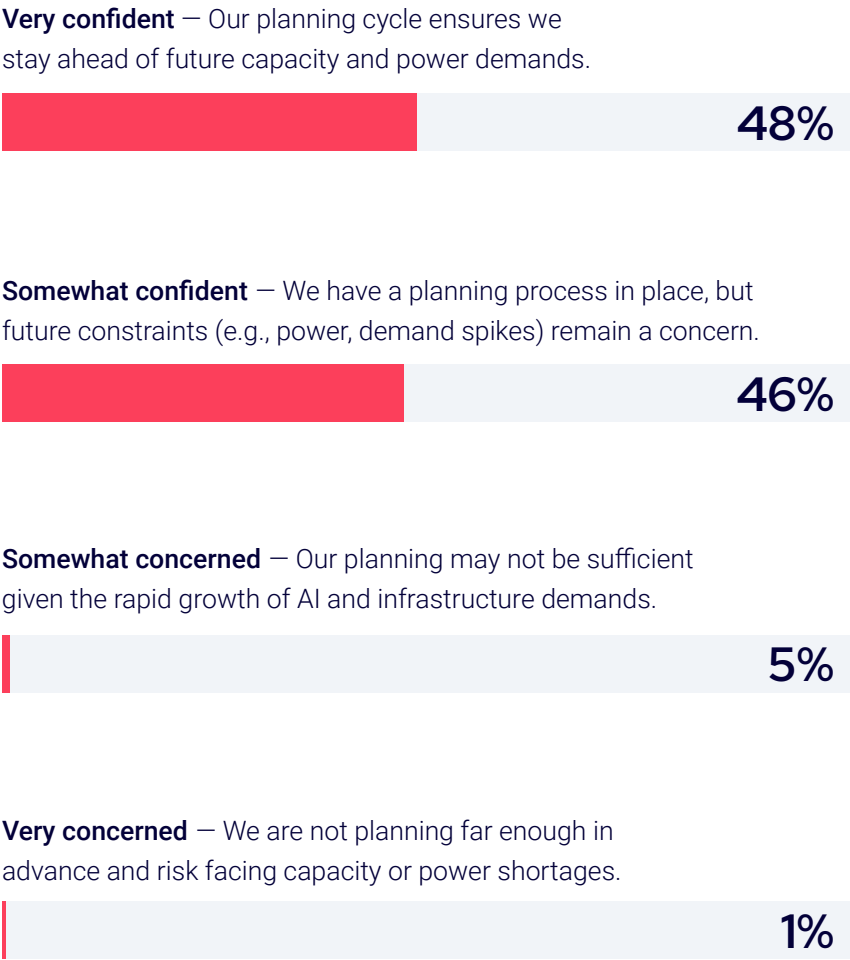*How far in advance is your organization planning its IT infrastructure and data center capacity needs?*

16%	62%	17%	6%
Less than one year	One year to three years	Three years to five years	More than five years

FIG. 16

How confident are you that your organization is planning far enough in advance for IT infrastructure and data center capacity needs?

Very confident — Our planning cycle ensures we stay ahead of future capacity and power demands.

48%

Somewhat confident — We have a planning process in place, but future constraints (e.g., power, demand spikes) remain a concern.

46%

Somewhat concerned — Our planning may not be sufficient given the rapid growth of AI and infrastructure demands.

5%

Very concerned — We are not planning far enough in advance and risk facing capacity or power shortages.

1%

Limited supply and rising costs

That low vacancy rate comes despite supply in primary data center markets increasing by 34%, with a record **6,350 MW under construction** in those markets. As of early last year, 80% of new construction had been pre-leased before facilities were operational.[18] The capacity available today is largely spoken for, and the remaining options **come at a premium**. Construction timelines are extending due to skilled labor shortages, supply chain constraints, and power availability challenges, all of which make last-minute procurement a high-risk strategy.[19]

AI-driven infrastructure is **intensifying** this strain. Global data center demand is projected to grow by as much as 22% annually through 2030, **requiring more than twice the capacity built since 2000** in just a fraction of the time.[20] Even if all known projects are delivered on schedule, the U.S. alone could face a data center supply shortfall exceeding **15 GW** by the start of the next decade. Racks that once supported **36 kW** workloads are now tasked with handling AI training models demanding **80 kW** or more, placing further pressure on power distribution and availability.

Procurement strategies must change

Traditional procurement strategies are proving **increasingly ineffective**. Securing power and space now requires committing earlier and at full capacity from Day 1 as providers move away from gradual deployment models. Organizations that wait until they need additional capacity may find themselves priced out or left without options altogether.

Companies that align **procurement strategies** with **market realities** will be best positioned to secure capacity as it becomes available. Successful enterprises have been **extending planning cycles to at least 24 months** and reassessing geographic preferences to expand their range of viable options.

As enterprises double down on AI's transformative potential, the competition for infrastructure will remain intense, leaving little room for hesitation.

Key takeaways

1

Because of mounting supply constraints, 79% of **organizations are beginning to plan their data center capacity needs** more than one year out. AI-driven demand is accelerating, and vacancy rates have hit historic lows.

2

Short-term planning leaves enterprises **vulnerable to infrastructure shortages**, yet a significant portion of organizations planning less than a year ahead still express high confidence in their readiness. Given the tightening market, securing data center capacity will require earlier commitments and more flexible procurement strategies.

SECTION 5: NETWORKING AND SECURITY CONCERNS

AI workloads require stronger networks and smarter security approaches

AI is reshaping how organizations safeguard their data and manage their processing power.

When asked to name where their **AI data is stored**, 60% of organizations reported using a private cloud, 48% use hybrid environments, and 47% use the public cloud [Fig. 17]. The percentage of organizations relying on GPU-as-a-service vendors increased from 34% last year to 40% today, while public cloud deployments grew from 30% to 34% [Fig. 18]. Meanwhile, the **public cloud remains the dominant choice** for storing AI training data, as 68% of organizations choose that option, and 54% turn to colocation [Fig. 19].

As infrastructure demands change, **companies are adapting** by blending private and public solutions to find the right balance of performance and cost.

AI workloads are straining data centers

AI workloads are putting **unprecedented pressure on data centers**. Bandwidth shortages and latency, once secondary concerns, have become major challenges. In the past year, **bandwidth shortages** affected 59% of respondents, up from 43% a year ago, while excessive latency rose from 32% to 53% [Fig. 20]. As AI use cases scale across industries, the need for lower-latency connectivity will only intensify.

Legacy networks are struggling to keep up as AI models move from **handling terabytes to petabytes** of data. Organizations using distributed AI architectures face the most pressure, as frequent data transfers between cloud, edge, and on-premises environments increase congestion and delays.

FIG. 17

Where is your AI data housed?

In a private cloud
60%

In a hybrid cloud environment
48%

In the public cloud
47%

In a colocation data center
20%

In an on-premises data center
20%

FIG. 18

Which of the following best describes where your organization deploys the most GPUs?

40%	Via a GPU-as-a-service vendor specializing in AI	**34%**	In the public cloud	**13%**	In our colocation data center	**13%**	In our on-premises data center

FIG. 19

Where is the data you utilize or plan to utilize for AI training or inference applications housed?

68% In the public cloud

54% In our colocation data center
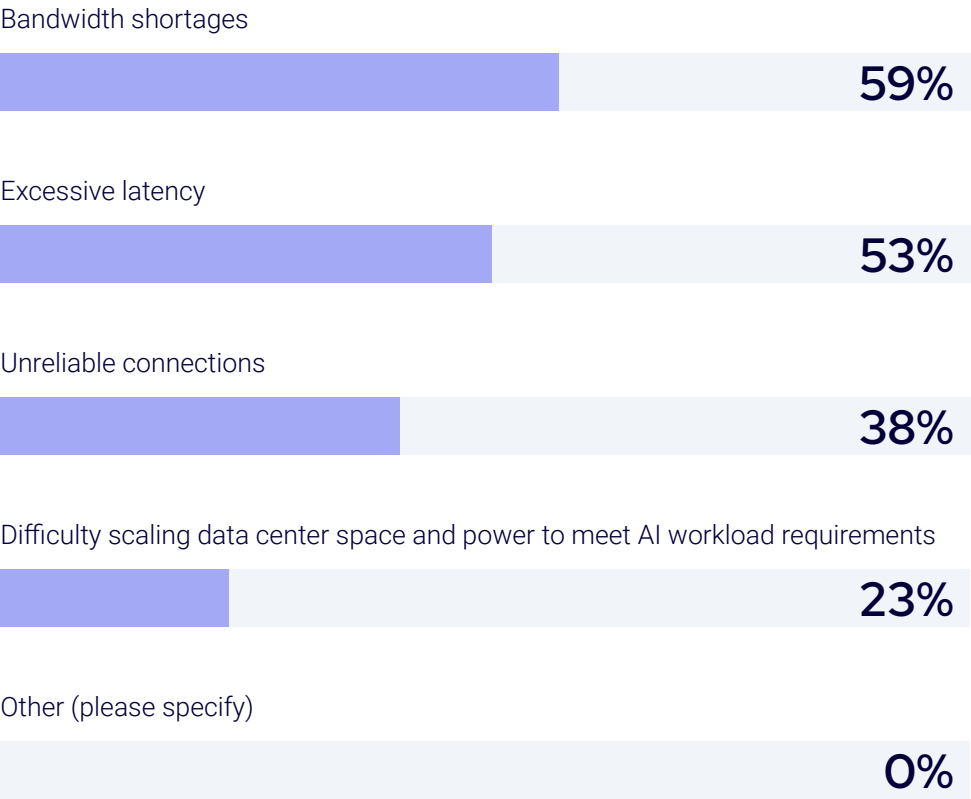
36% In our on-premises data center
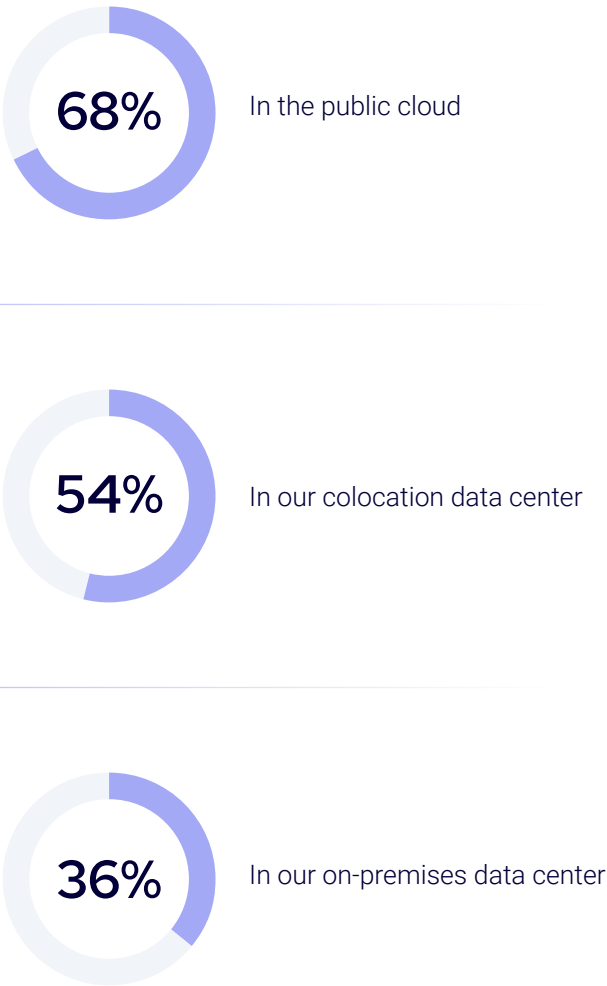
Scaling AI requires more than space and power

Despite connectivity challenges, **some infrastructure constraints have eased**. Reports of difficulty scaling data center space and power dropped from 34% to 23%, **indicating progress** in expanding capacity. [Fig. 20].

However, **solving power and space constraints won't ease performance issues if networks remain a bottleneck**. As AI adoption accelerates, enterprises must rethink their network strategies—prioritizing proximity, high-speed interconnects, and redundancy—to support next-generation applications.

FIG. 20

In the past 12 months, have you encountered any of the following performance issues with your AI applications or workloads?

Bandwidth shortages

59%

Excessive latency

53%

Unreliable connections

38%

Difficulty scaling data center space and power to meet AI workload requirements

23%

Other (please specify)

0%

Which of the following tactics is your organization implementing to reduce performance issues for its AI applications or workloads?

Using 5G networks (e.g., 5G-enabled IoT devices, etc.)

65%

Implementing software-defined networking (SDN)

55%

Using network function virtualization (NFV)

51%

Using Wi-Fi 6 or Wi-Fi 7

48%

Using third-party colocation data centers to process data closer to the edge of the network

37%

Deploying AI hardware on-premises

32%

None of the above

0%

Networks must evolve to keep pace with AI

AI's success depends on **network performance**, and organizations are taking deliberate steps to **improve performance** and **reduce latency** as AI workloads grow in complexity.

In the past year, the use of **5G networks to enhance AI performance** has jumped from 54% to 65%. Software-defined networking (SDN) adoption climbed from 38% to 55%, while network function virtualization (NFV) rose from 45% to 51% [Fig. 21]. These shifts indicate a growing realization that traditional networking is insufficient for AI's bandwidth and latency demands.
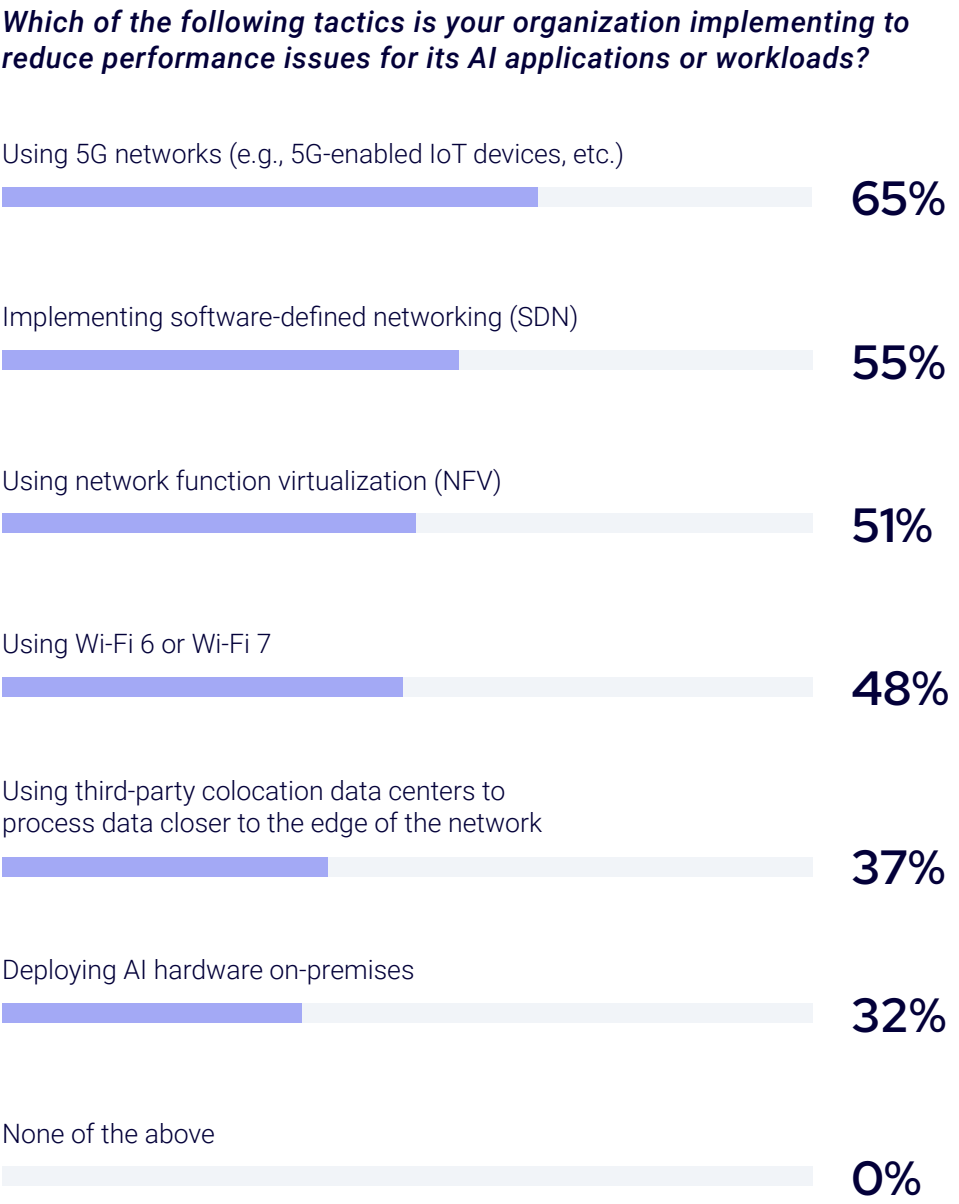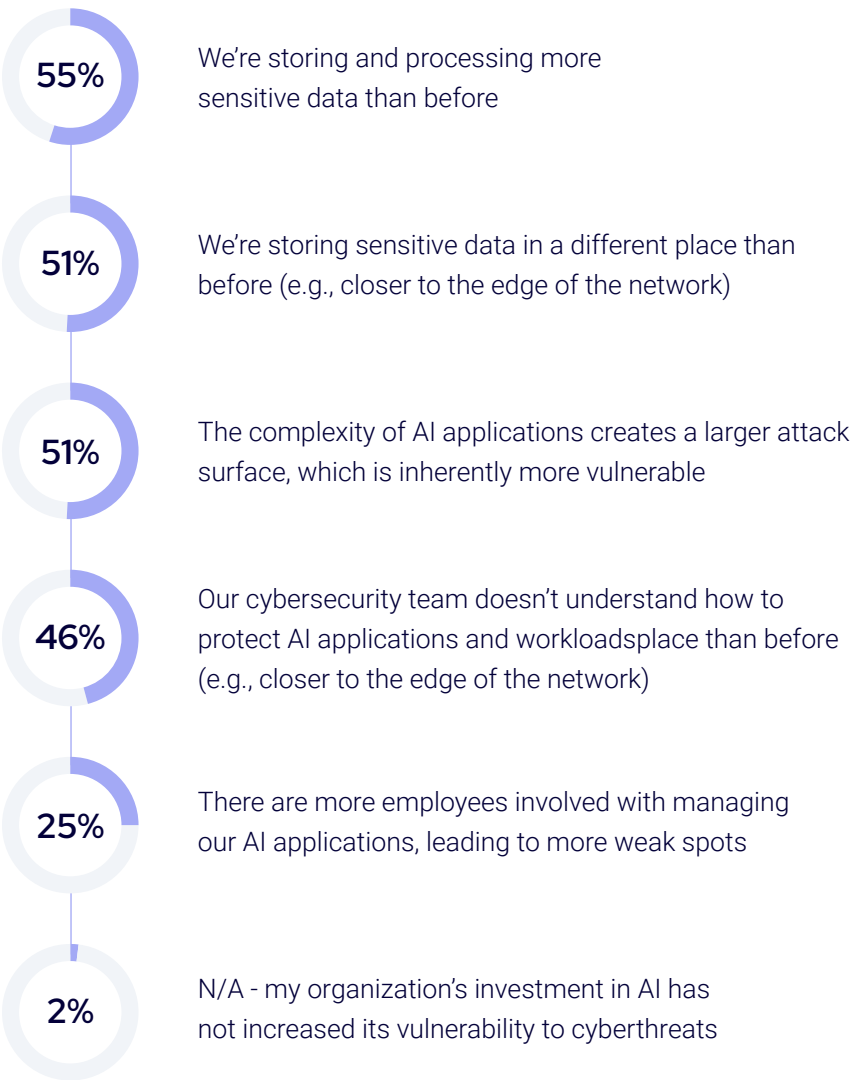
FIG. 22

How has increasing your organization's investment in AI increased its vulnerability to cyberthreats?

55% We're storing and processing more sensitive data than before

51% We're storing sensitive data in a different place than before (e.g., closer to the edge of the network)

51% The complexity of AI applications creates a larger attack surface, which is inherently more vulnerable

46% Our cybersecurity team doesn't understand how to protect AI applications and workloadsplace than before (e.g., closer to the edge of the network)

25% There are more employees involved with managing our AI applications, leading to more weak spots

2% N/A - my organization's investment in AI has not increased its vulnerability to cyberthreats

AI growth is creating new cybersecurity gaps

As AI-driven systems process vast amounts of sensitive data, **security challenges are mounting**. This year, 55% of respondents said AI adoption has increased their vulnerability to **cyber threats** because they're storing and processing more sensitive data than before—up from 39% last year [Fig. 22]. AI applications often integrate data from multiple sources, **exposing security gaps** that legacy frameworks weren't built to handle.

Security teams are also **struggling to keep pace**. This year, 46% of respondents reported that their cybersecurity teams lack a full understanding of AI security—up six points from last year. Meanwhile, 51% said AI inherently expands the attack surface [Fig. 22]. As AI adoption grows, **strengthening cybersecurity frameworks will be key** to minimizing risk and protecting AI-driven data.

Key takeaways

1

Organizations are **expanding their AI infrastructure** by adopting a mix of private, hybrid, and public cloud solutions. This approach reflects the need for flexible, scalable environments that balance performance, security, and cost as AI workloads grow.

2

AI adoption is accelerating, but security strategies are struggling to keep pace, leaving **organizations more exposed to cyber risks than before**. Companies must adapt their cybersecurity frameworks to address AI's unique vulnerabilities as sensitive data volumes grow and attack surfaces expand.

SECTION 6: SUSTAINABILITY PRESSURE

Businesses face growing pressure to adopt sustainable infrastructure

Organizations continue to prioritize sustainability in their IT infrastructure despite easing regulatory pressures.

The recent reduction in Environmental Protection Agency (EPA) funding comes from an overarching belief that **strict ESG mandates could hinder economic growth**. Yet, market forces remain a stronger influence than regulation.

Institutional investors—including those located outside the U.S.—assess companies based on **sustainability commitments**, and many major funds integrate ESG criteria into their decisions. Meanwhile, consumers and business partners increasingly weigh environmental impact when choosing whom to support.

FIG. 23

Are you willing to pay higher costs for your data centers or third-party cloud vendors to use clean or renewable energy and/or buy credits to offset their carbon footprints?

Yes, significantly more (more than 20% price increase)

27%

Yes, somewhat more (11-20% price increase)

51%

Yes, slightly more (1-10% price increase)

21%

No, I am not willing to pay more

1%

Sustainability pressures persist

When respondents think back to where they were a year ago, 79% feel there's **increased pressure** to make their infrastructure more sustainable [Fig. 24]. More than a quarter (27%) are willing to pay at least **20% higher costs** to ensure their data centers or cloud providers use **renewable energy** or **purchase carbon offsets**, and roughly half (51%) would pay 11-20% more [Fig. 23].

These findings suggest that sustainability is **no longer just a compliance issue**. For the organizations surveyed, sustainability has become an operational and strategic priority, shaped by investor expectations, consumer preferences, and long-term cost efficiencies.

AI driving energy consumption to new heights

The rising costs of IT infrastructure are closely tied to increasing energy demand, **particularly from AI workloads**. Data centers accounted for 4.4% of total U.S. electricity consumption in 2023, and that figure is projected to rise to as much as 12% by 2028.[21]

The demand for AI-driven computing power, such as high-performance GPUs developed by companies like NVIDIA, is driving a **surge in energy consumption**. Average power densities have more than doubled in the last two years, from **8 kW per rack to 17**, with **projections reaching 30 kW** per rack by 2027.[22] This trend is compounded by AI models, which could require more than **5 GW** of electricity by 2030—the equivalent of Manhattan's entire energy demand at any given time.[23] Scaled GPU-centric deployments will see an average of 100 kW per rack in the coming years.

FIG. 24

Compared to a year ago, how much pressure do you feel to make IT infrastructure more sustainable?

28% It has increased significantly

51% It has increased somewhat

15% It has remained about the same

5% It has decreased somewhat

1% It has decreased significantly

1% My organization did not feel any pressure to make IT infrastructure more sustainable

The hidden cost of cooling

Water usage is another growing concern. Hyperscale data centers consumed **66 billion liters of water in 2023**,[24] and projections suggest this figure **could rise to as much as 275 billion liters by 2028** without more efficient cooling strategies. To put this in perspective, 66 billion liters would fill 26,400 Olympic-sized swimming pools—enough to cover Leesburg, Virginia, just outside "data center alley." At 275 billion liters, that number jumps to 110,000 pools, an area the size of Anaheim, California.

In response, companies are investing in more efficient cooling technologies—including **chip manufacturers' shift to liquid cooling**, which circulates coolant through tubes rather than relying on air, and emerging methods like **submerging racks in non-conductive liquid to dissipate hea**t. Microsoft, for example, has worked to reduce water intensity in its data centers by 80% since the early 2000s,[25] while in Finland, Google's seawater cooling approach supplies excess heat to local homes.[26]

Search for reliable, sustainable energy sources

Despite these steps, **energy sourcing remains a pressing challenge**. Renewable energy is a preferred solution, but the intermittent nature of wind and solar generation presents reliability issues because data centers require 24/7 availability. Companies are addressing this challenge through a combination of long-term power purchase agreements, small modular nuclear reactors (SMRs), and battery storage systems.

However, supply remains constrained. In states like Virginia and Oregon, where **data centers could account for up to 46% and 24% of total electricity demand by 2030**, respectively, local governments are reassessing how to balance energy availability with economic growth.[27] Interestingly, the same AI technologies that are fueling increased energy use may also be deployed to manage it.

These pressures highlight why **many businesses are willing to absorb higher costs to maintain sustainable practices**. Regulatory enforcement may fluctuate, but organizations recognize that sustainability is a necessity for risk management, long-term stability, and investor confidence.

Key takeaways

1

Sustainability remains a rising priority for IT leaders despite shifting federal policies. Nearly four out of five respondents **feel more pressure to make their infrastructure sustainable** than they did a year ago, reinforcing that market and operational forces are driving change independently of regulation.

2

Businesses signal they are willing to **absorb higher costs to ensure their data centers use clean energy**. More than three-quarters of respondents said they would pay at least 11% more for renewable energy or carbon offsets.

CONCLUSION

It's *already* time for a fresh approach to AI infrastructure

AI is reshaping how organizations operate and innovate, requiring IT leaders to rethink their infrastructure strategies. To stay ahead, they must:

- ⦿ **Adopt scalable, high-density compute solutions** that accommodate AI's growing demands while seamlessly integrating emerging technologies.
- ⦿ **Utilize software-driven interconnection** to enable seamless collaboration and efficient management of distributed workloads and enable low latency connectivity to AI data sources.
- ⦿ **Integrate advanced cooling solutions**, such as liquid cooling, to meet sustainability objectives without compromising system performance.
- ⦿ **Implement comprehensive data management and end-to-end security measures** that safeguard sensitive AI-driven data and align with shifting regulatory requirements.
- ⦿ **Plan for capacity** by forecasting AI workload growth and securing access to colocation or cloud resources before demand exceeds availability.

Whether you're launching new products or sustaining your competitive advantage, your success hinges on making these foundational infrastructure investments. Partnering with a seasoned data center provider ensures access to the low-latency connections, high bandwidth, and optimized performance necessary to support your AI-driven growth.

Is your AI infrastructure built for what's next?

Flexential can help you advance your AI strategy.

Partner with us to build a scalable, sustainable AI infrastructure that supports your organization's growth and innovation.

Schedule a Consultation Today

FLEXENTIAL

Methodology

In February 2025, Flexential surveyed 355 IT decision-makers at the director level or above at organizations with over $100 million in annual revenue. All respondents had knowledge of their organizations' AI implementation and related infrastructure buildouts. Respondents came from a range of industries.

ANNUAL REVENUE

$101 million to $500 million	43%
$501 million to $2 billion	29%
More than $2 billion	28%

JOB LEVEL

Director	40%
Vice President	34%
C-suite	26%

INDUSTRY

Technology/IT Services/Software	37%
Financial Services	16%
Manufacturing	14%
Retail/Wholesale	11%
Telecom/Networks	8%
Healthcare	4%
Construction	3%
Entertainment/Media	2%
Business and Professional Services/B2B Services	1%
Education	1%
Food Service/Hospitality	1%
Public Sector/Government	1%
Transportation/Logistics	1%
Utilities	1%

References

1 "2024 Business Opportunity of AI." IDC, November 2024.
https://info.microsoft.com/ww-landing-business-opportunity-of-ai.html.

2 Vien, Courtney. "How Walmart's seen ROI on gen AI." CFO Brew, August 23, 2024.
https://www.cfobrew.com/stories/2024/08/23/how-walmart-s-seen-roi-on-gen-ai.

3 O'Brien, Matt. "What Is DeepSeek, the Chinese AI Company Upending the Stock Market?"
Associated Press, January 28, 2025.
https://apnews.com/article/deepseek-ai-china-f4908eaca221d601e31e7e3368778030.

4 Criddle, Cristina and Eleanor Olcott. "OpenAI says it has evidence China's
DeepSeek used its model to train competitor." Financial Times, January 28, 2025.
https://www.ft.com/content/a0dfedd1-5255-4fa9-8ccc-1fe01de87ea6.

5 Mullen, Caitlin. "Morgan Stanley CEO outlines AI benefits." Banking Dive, June 11, 2024.
https://www.bankingdive.com/news/morgan-stanley-ceo-ted-pick-ai-banking-investment-wealth-dan-simkowitz/718571/.

6 "Ma'aden empowers employees with Microsoft 365 Copilot and Azure
OpenAI Service, saving 2,200 hours monthly." Microsoft, January 29, 2025.
https://www.microsoft.com/en/customers/story/21514-maaden-microsoft-365-copilot.

7 "New Study from Gregory FCA Finds That AI-Enablement Increases Firmwide Productivity
10.2% While Spurring Client and Employee Retention." Gregory FCA, August 13, 2024.
https://gregoryfca.com/insights/ai-increases-firmwide-productivity/.

8 "Acentra Health boosts employee productivity with generative AI and Azure OpenAI Service,
saving 11,000 nursing hours and nearly $800,000." Microsoft, December 3, 2024.
https://www.microsoft.com/en/customers/story/19280-acentra-health-azure.

9 "2024 Business Opportunity of AI." IDC, November 2024.
https://info.microsoft.com/ww-landing-business-opportunity-of-ai.html.

10 "YouGov Survey: AI and Jobs." YouGov, August 23, 2024.
https://d3nkl3psvxxpe9.cloudfront.net/documents/AI_and_Jobs_poll_results_u47r4j0.pdf.

11 "YouGov Survey: Artificial Intelligence." YouGov, December 12, 2024.
https://d3nkl3psvxxpe9.cloudfront.net/documents/Artificial_Intelligence_poll_results.pdf.

12 "AI: The Ambitions Are Bold, but the Talent Is Scarce." Bain, February 24, 2025.
https://www.bain.com/insights/ai-the-ambitions-are-bold-but-the-talent-is-scarce-snap-chart/.

13 Rattner, Nate. "How the AI Talent Race Is Reshaping the Tech Job Market."
Wall Street Journal, March 10, 2025.
https://www.wsj.com/tech/ai/how-the-ai-talent-race-is-reshaping-the-tech-job-market-93df0615.

14 "AI: The Ambitions Are Bold, but the Talent Is Scarce." Bain, February 24, 2025.
https://www.bain.com/insights/ai-the-ambitions-are-bold-but-the-talent-is-scarce-snap-chart/.

15 "AI skills gap widens: 71% of AI talent are men, while only 22% of baby boomers receive training."
Randstad, November 12, 2024.
https://www.randstad.com/press/2024/ai-skills-gap-widens/.

16 "North America Data Center Trends H2 2024." CBRE, February 26, 2025.
https://www.cbre.com/insights/reports/north-america-data-center-trends-h2-2024.

17 Batson, Andrew. "North America Data Center Report Year-end 2024." JLL, March 9, 2025.
https://www.jll.com/en-us/insights/market-dynamics/north-america-data-centers.

18 "North America Data Center Trends H1 2024." CBRE, August 19, 2024.
https://www.cbre.com/insights/reports/north-america-data-center-trends-h1-2024.

19 Salzman, Avi. "Energy Execs Warn Electricians Are Hard to Find. It Could Slow the AI Data Center Boom."
Barron's, March 14, 2025.
https://www.barrons.com/articles/electricians-ai-shortage-blackrock-microsoft-db2fe5bc.

20 "AI power: Expanding data center capacity to meet growing demand." McKinsey, October 29, 2024.
https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/ai-power-expanding-data-center-capacity-to-meet-growing-demand.

21 "2024 United States Data Center Energy Usage Report." Lawrence Berkeley National Laboratory,
December 19, 2024.
https://eta.lbl.gov/publications/2024-lbnl-data-center-energy-usage-report.

22 "AI power: Expanding data center capacity to meet growing demand." McKinsey, October 29, 2024.
https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/ai-power-expanding-data-center-capacity-to-meet-growing-demand.

23 Hiller, Jennifer. "Five Things to Know About AI's Thirst for Energy." Wall Street Journal, February 1, 2025.
https://www.wsj.com/tech/ai/ai-energy-electricity-use-what-to-know-8c9e64b7.

24 "2024 United States Data Center Energy Usage Report."
Lawrence Berkeley National Laboratory, December 19, 2024.
https://eta.lbl.gov/publications/2024-lbnl-data-center-energy-usage-report.

25 Walsh, Noelle. "Sustainable by design: Transforming datacenter water efficiency."
Microsoft, July 25, 2024.
https://www.microsoft.com/en-us/microsoft-cloud/blog/2024/07/25/sustainable-by-design-transforming-datacenter-water-efficiency/.

26 "Google's first-ever heat recovery project for neighbourhoods in Finland."
Posted May 20, 2024, by Google. YouTube, 3 min., 04 sec.
https://www.youtube.com/watch?v=YUmFQM4jvD8.

27 "Powering Intelligence: Analyzing Artificial Intelligence and Data Center Energy Consumption."
Electric Power Research Institute, May 2024.
https://www.epri.com/research/products/3002028905.