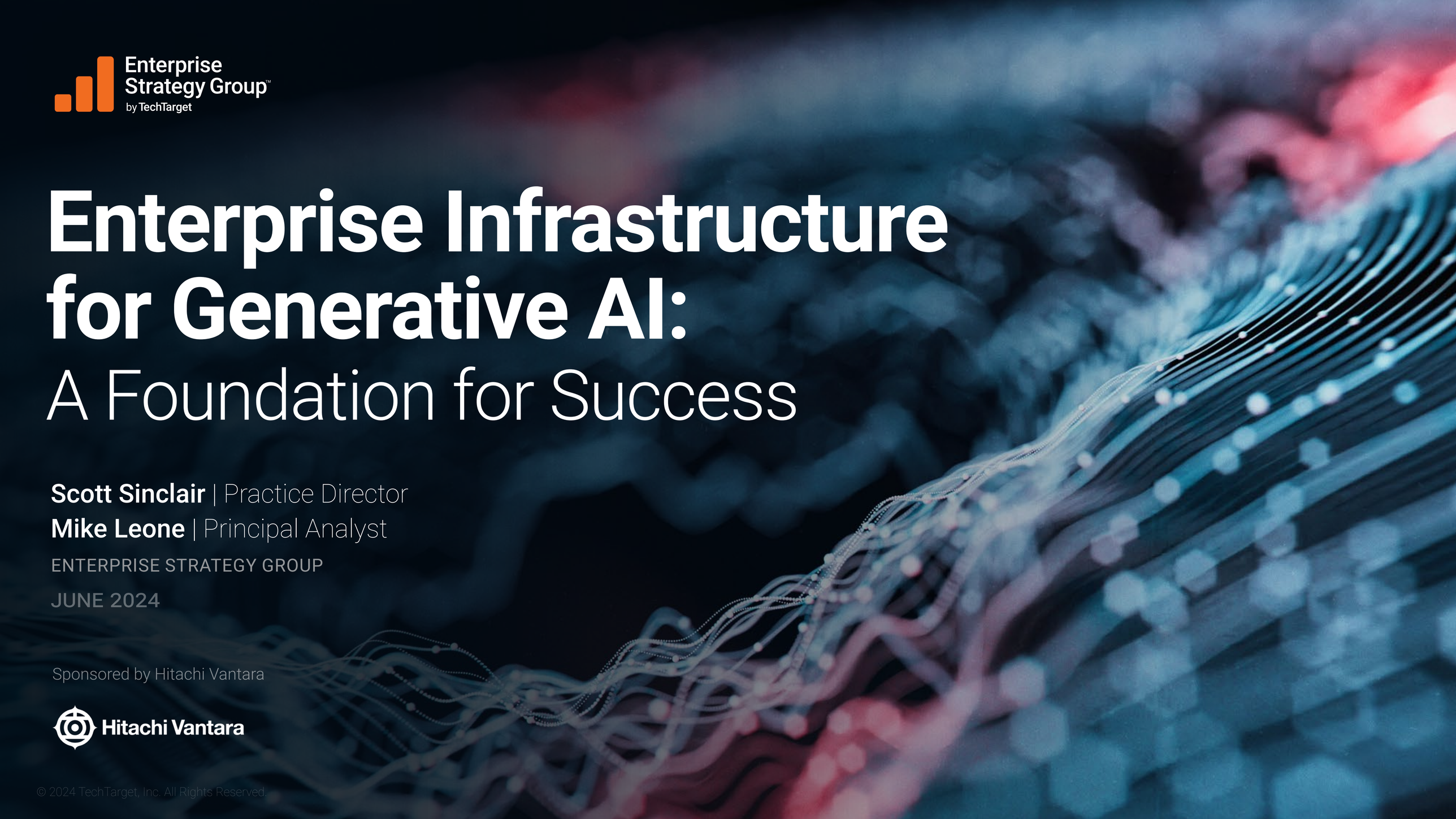


Enterprise Strategy Group™ by TechTarget

# Enterprise Infrastructure for Generative AI:

## A Foundation for Success

**Scott Sinclair** | Practice Director
**Mike Leone** | Principal Analyst
ENTERPRISE STRATEGY GROUP

JUNE 2024

Sponsored by Hitachi Vantara

Hitachi Vantara

# Introduction and Research Objectives

Generative AI, or GenAI, represents a transformative technology with the potential to revolutionize how businesses operate and compete. However, unlocking the true power of GenAI requires a strong foundation—robust and secure infrastructure that can handle the demands of this powerful technology. This report delves into the critical role of data infrastructure for enterprise GenAI success. Beyond examining the technology landscape and key trends, we'll place particular emphasis on the decisions surrounding data, models, and most importantly, the infrastructure that underpins successful GenAI implementations.

**By reading this eBook, you will:**

- Gain a comprehensive understanding of the role of enterprise data infrastructure in successful GenAI adoption.
- Learn about key trends and challenges driving GenAI within organizations.
- Discover best practices for selecting data, models, and building a secure and scalable infrastructure.
- Explore various use cases and the potential benefits of GenAI for your enterprise.
- Gain insights into the technical considerations and vendor landscape to make informed decisions.

# Audience and Highlighted Findings

## Audience

The target respondents for this survey consisted of a mix of enterprise IT leaders, line-of-business leaders, and data-centric practitioners/management (e.g., data engineering, data science personas either in dedicated data science teams or embedded in lines of business). This enabled the research to provide a comprehensive view of GenAI initiatives from strategy and decision-making, to infrastructure and implementation, to operations and ongoing management in production.

The respondents represented organizations based in North America (U.S., CA) and Western Europe (U.K., DE, FR) across all public- and private-sector industries. This includes minimum targets of up to 10% within key industries, including financial services, tech/telco/media and entertainment, manufacturing, and healthcare/life sciences.

## Highlighted Findings

63%

of organizations have significant room to improve their enterprise infrastructure and data ecosystem preparedness to support GenAI initiatives.

78%

of respondents cite some mix of on-premises and public cloud for building and using GenAI solutions.

96%

of organizations prefer an alternative to proprietary models, driving the open source model movement.

86%

of organizations will lean on retrieval-augmented generation to some extent to incorporate enterprise data.

56%

view improving accuracy as the top reason for retraining a GenAI model.

53%

believe their IT teams are unable to keep up with the pace of innovation being driven by GenAI.

61%

believe there is significant room to improve collaboration among teams responsible for selecting, implementing, and managing GenAI solutions and infrastructure.

**Security, cost, and data quality**

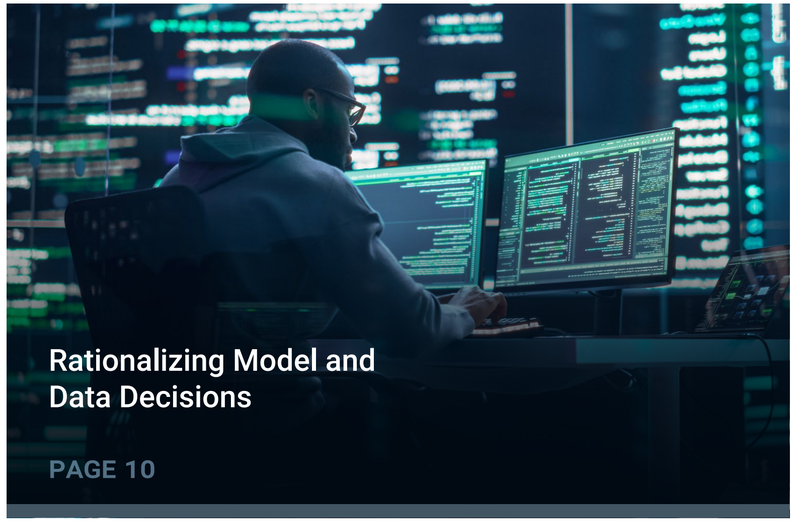
are some of the biggest challenges and concerns organizations face when it comes to storing and managing data related to GenAI initiatives.



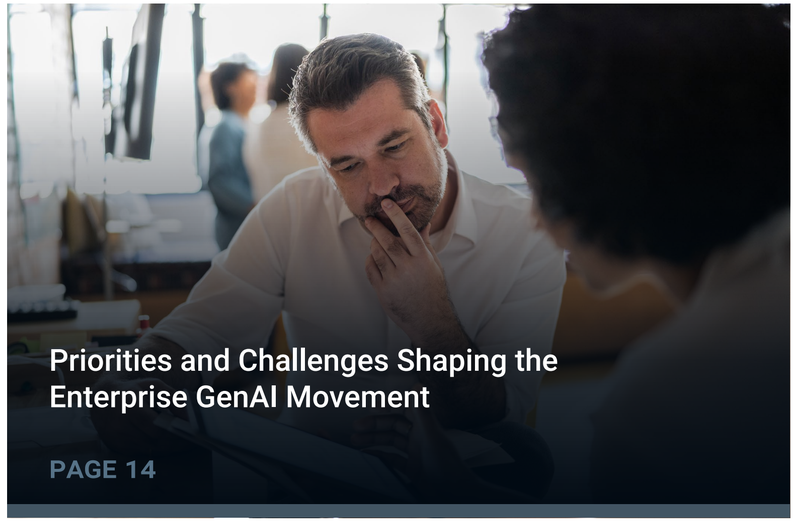
BACK TO CONTENTS

# Contents





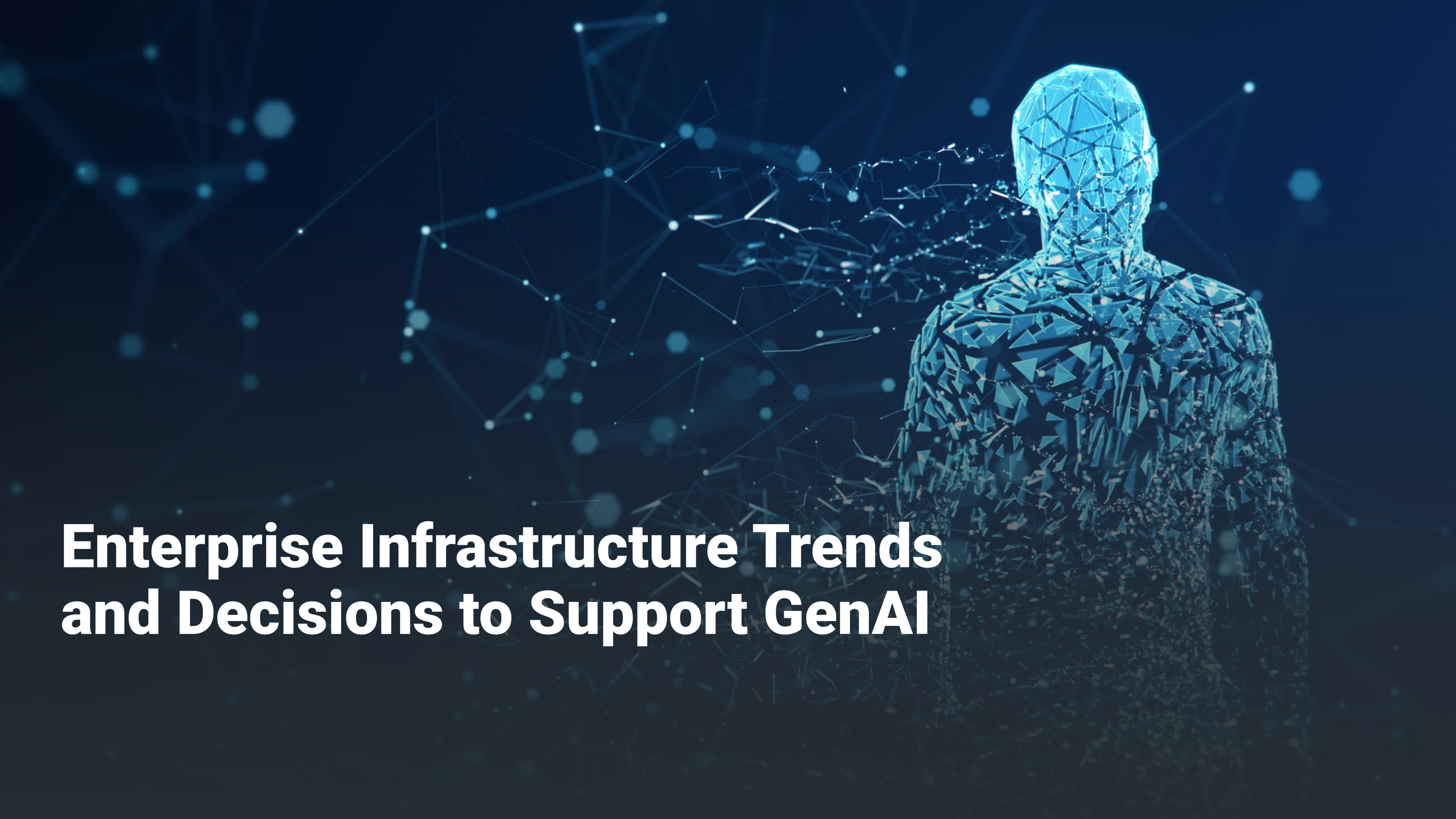












# Enterprise Infrastructure Trends and Decisions to Support GenAI

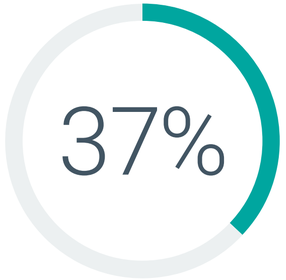
# A Balanced Approach Is Preferred When It Comes to the Underlying Infrastructure Used for GenAI

Organizations believe their infrastructure is in good shape to support GenAI initiatives. This belief might shift, however, as GenAI adoption maturity increases.

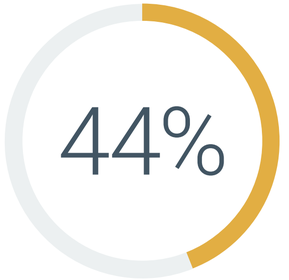
It's clear that hybrid cloud is essential, with most organizations using a mix of on-premises and public cloud to support GenAI deployments. Of those surveyed, 78% of respondents cited some mix of on-premises and public cloud services for building and using GenAI solutions.

The desire for hybrid cloud approaches extends to data pipelines: Underlying data pipelines help move and manage data based on the use case or initiative at hand.

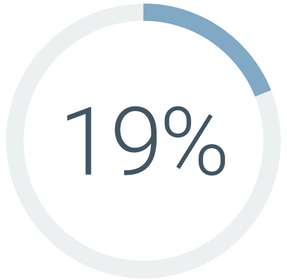
## Self-assessing Infrastructure Readiness




**Highly prepared** - our infrastructure and data ecosystem are well-suited for implementing generative AI solutions

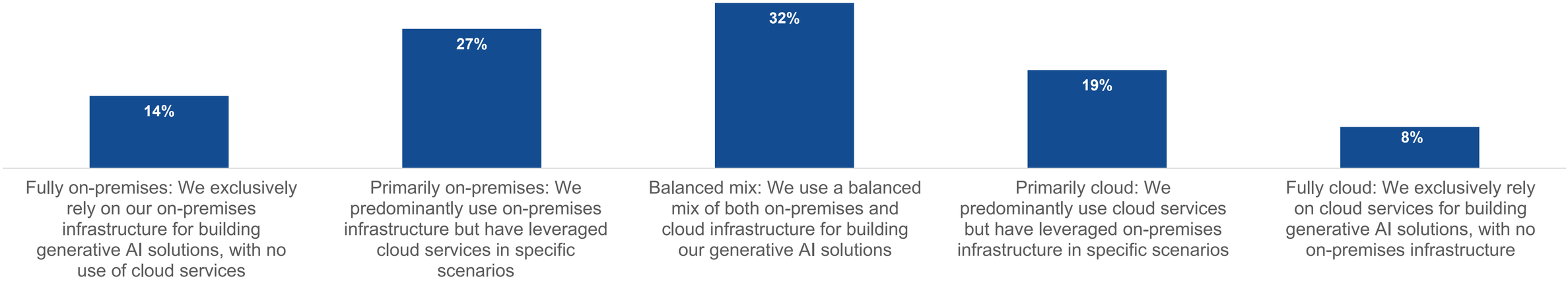



**Moderately prepared** - we have some capabilities but may require enhancements for implementing generative AI solutions




**Somewhat prepared** - we have a few capabilities but may require several enhancements for implementing generative AI solutions

## Infrastructure Preferences When Deploying GenAI Solutions






BACK TO CONTENTS

# Security, Quality, and Flexibility Lead Critical Infrastructure Characteristics

As we think about the whole supporting data and IT infrastructure, several characteristics emerge as essential:

**1.**

Robust security measures to protect sensitive data and models from breaches and misuse.

**2.**

High-quality data, integral for training reliable and unbiased GenAI models.

**3.**

Flexible data infrastructure components, which best accommodate diverse data sources, formats, and evolving requirements, and enable seamless integration and scalability as GenAI use cases expand.

**The Most Important Characteristics in Enterprise Data and Storage Infrastructure**

37%

Data security

31%

The ability to ensure data quality

28%

Flexible and adaptable to changing data demands

28%

Scalability to accommodate growing data volume

28%

Ability to track data and its usage/performance

28%

Integration across modern and legacy systems

28%

Cost management

28%

Data backup and recovery

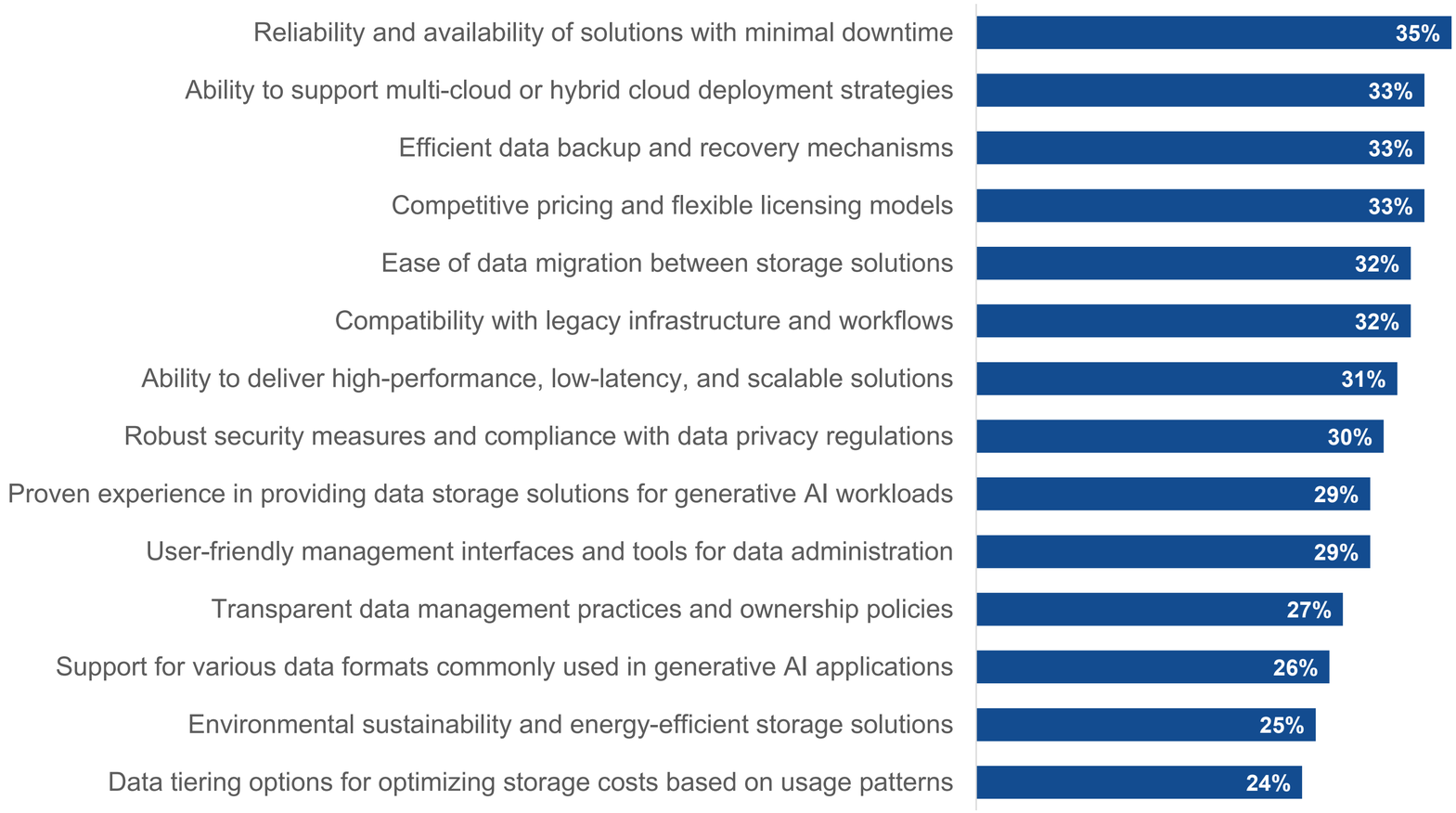


BACK TO CONTENTS

## Storage Vendor Solutions Must Be Comprehensive

While organizations consider the underlying data and storage requirements to support GenAI, they also focus on availability, hybrid cloud support, efficient data backup options, and pricing flexibility. On average, organizations identify more than four criteria from this list of requirements that should serve as a checklist for organizations when evaluating and selecting GenAI infrastructure vendors who can deliver the robust support and seamless integration necessary to deliver enterprise-ready AI.

### Data Storage Vendor Selection Criteria for GenAI Apps






BACK TO CONTENTS

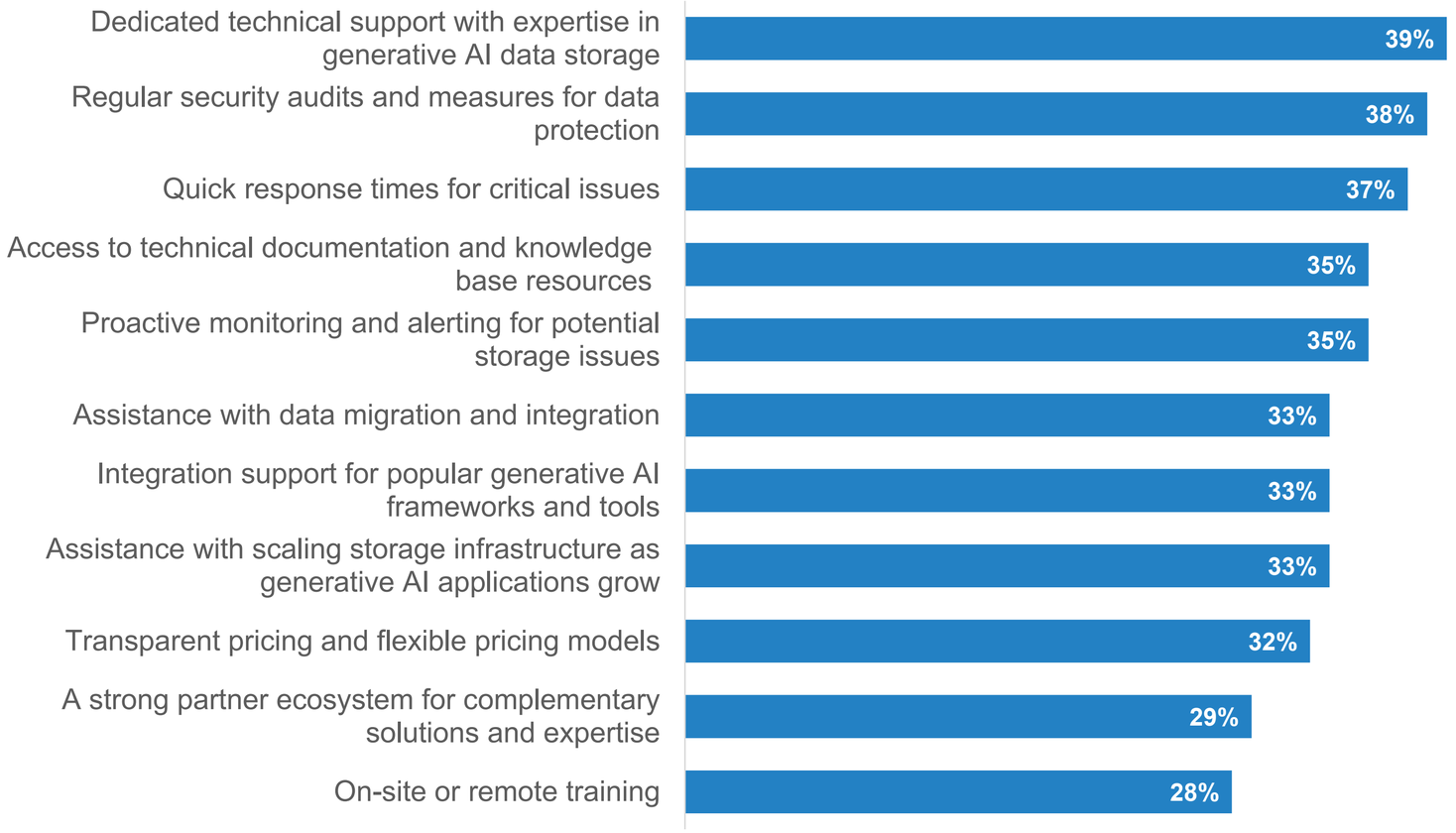
"Technical criteria, such as fast performance and low latency, **are all but expected at this point."**

## Storage Vendors With Deep GenAI Knowledge, Security Chops, and Fast Response Times Underpin a Long List of Expectations

When selecting a data storage vendor for GenAI, organizations expect a range of support and expertise to ensure seamless integration and optimal performance. Technical criteria, such as fast performance and low latency, are all but expected at this point.

Given the number of skills gaps that exist today within every organization, the ideal data storage vendor should also prioritize security and offer comprehensive training and support programs. These programs can help organizations better understand the technology, rapidly gain value, and, ultimately, maximize their return on investment.

**GenAI Support Expected From a Storage Vendor**






BACK TO CONTENTS

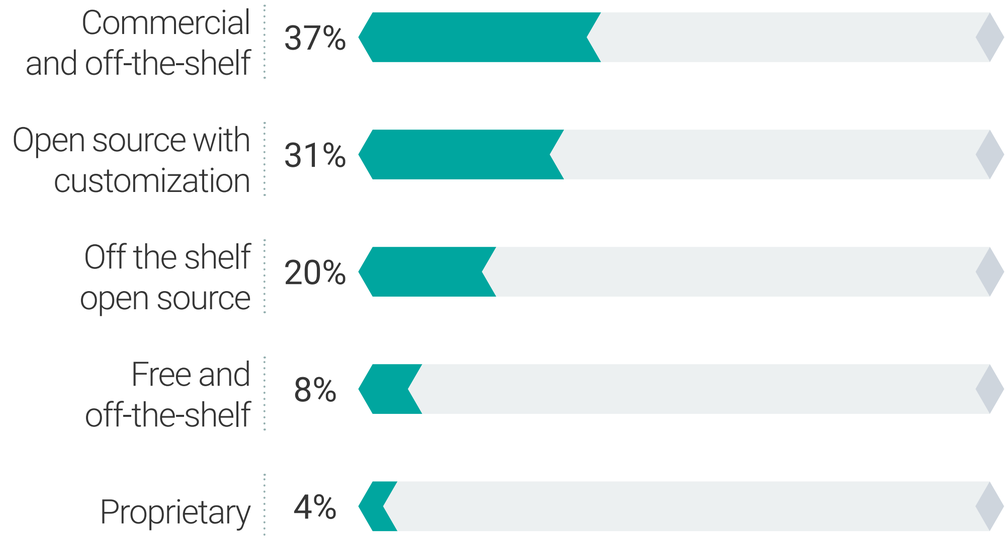
# Rationalizing Model and Data Decisions

# Organizations Are Using a More Generic and Affordable Approach to Employing GenAI Technology Today

With 96% of organizations preferring an alternative to proprietary large language models (LLMs), organizations are actively seeking out lower-cost options to get started with GenAI.

Over the long term, however, organizations expect the use of proprietary models to increase as they gain expertise—likely in an effort to achieve competitive differentiation.

### Current LLM Preferences




### Long-term LLM Preferences


**Mostly proprietary:** We prefer to predominantly use proprietary (internally developed) tools and technologies with some external elements.


**Mostly open source:** We prefer to predominantly use open source tools and technologies with some reliance on commercial and/or proprietary components.


**Mostly commercial:** We prefer to predominantly use commercial tools and technologies (developed by third-party companies) with some open source and/or proprietary elements.


**Balanced mix:** We prefer a balanced mix of commercial, proprietary and open source tools and technologies in building our generative AI solutions.



BACK TO CONTENTS

# Incorporating Enterprise Data Is Proving to Be the Business Differentiator for GenAI

As organizations seek to customize models and incorporate enterprise data and knowledge bases, **there are two dimensions to consider:**

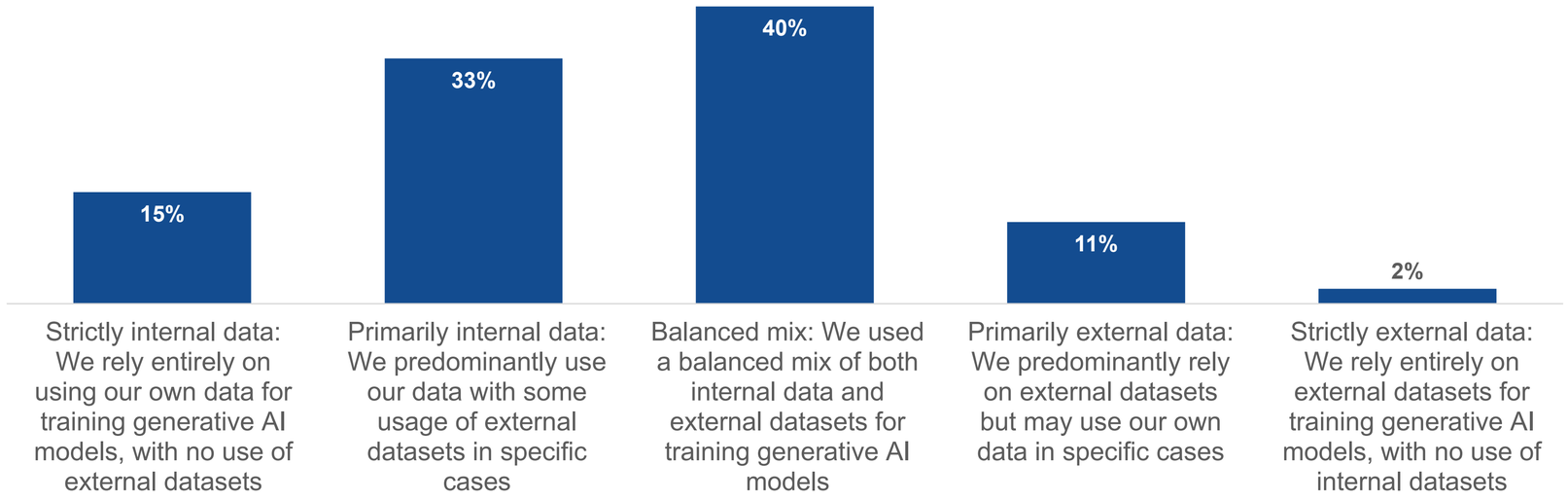
**1.**

What data is being used: Of organizations surveyed, **84% cited some balance of internal and external data to be used for training models.** This is especially true among IT respondents.

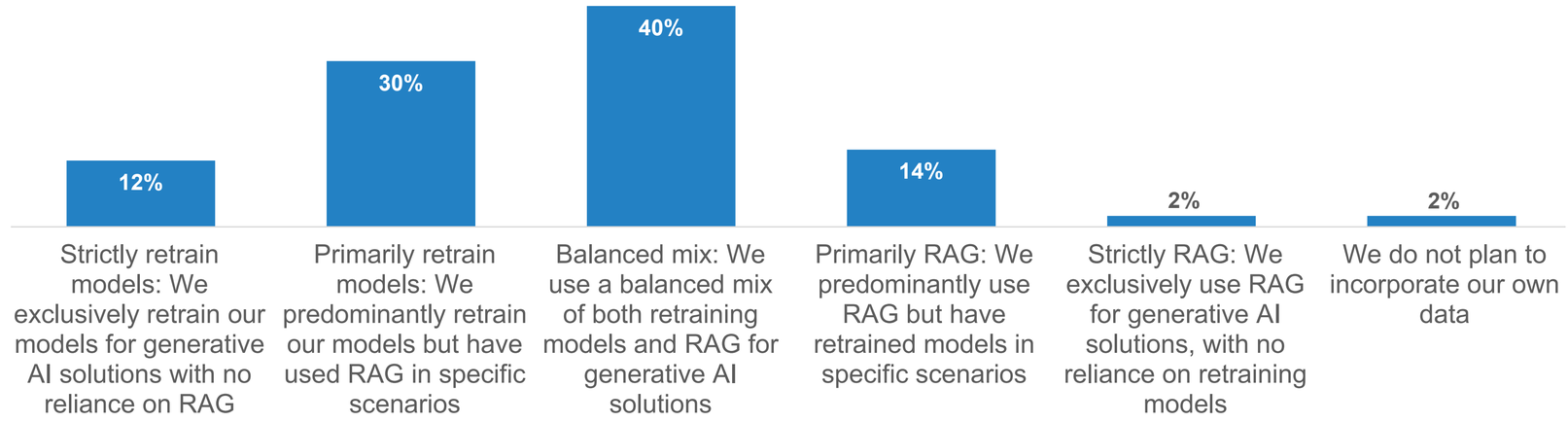
**2.**

The approach to incorporating data: train/retrain/fine-tune vs. retrieval-augmented generation (RAG). **Our research showed 86% of organizations will lean on RAG to some extent.**

## Data Used for Training GenAI Models




## Strategy for Incorporating Data Into GenAI Models






BACK TO CONTENTS

# Accuracy Remains King as the Top Reason for Retraining GenAI Models

While these are new techniques, there are pros and cons to the different methods—mostly around cost, control, and timeliness of data.

The desire for improved accuracy tops the list of reasons for ensuring the most relevant and recent data gets incorporated into an LLM, followed by the desire to keep pace with technology, regulations, and shifting data patterns. Managing data will not only enable greater levels of accuracy, but it will also help to maintain accuracy as data and business conditions evolve.

**Reasons for Retraining an LLM**

- 56% — Improve accuracy
- 47% — Keep pace with technological advancements
- 43% — Address compliance and/or regulatory requirements
- 42% — Adapt for data draft/ changing data patterns
- 41% — Respond to feedback of end-users
- 40% — Adapt to a change in the use case
- 32% — Address bias/ fairness concerns



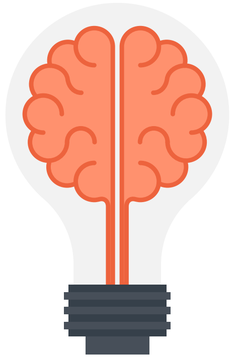
BACK TO CONTENTS

# Priorities and Challenges Shaping the Enterprise GenAI Movement

## Companies Remain Steadfast in Pursuing GenAI

Despite the relative infancy of the GenAI space, the technology is being embraced universally. Most organizations are already leveraging GenAI technology for at least one use case, whether independently or organizationally managed.

In addition, 90% of organizations believe investing in GenAI will enhance business operations by improving operational efficiency and employee productivity. It's also expected that those organizations investing in GenAI early and often are likely to gain a competitive advantage over their laggard peers.



97%

of organizations view GenAI initiatives **as a top-five priority.**

### GenAI Market Penetration

30%

We are leveraging the technology for several use cases across our businesses.

22%

We are leveraging the technology for our initial use case.

11%

We have definitive plans and have identified at least one use case.

33%

We are interested but are still exploring use cases.

3%

We have no plans or interest.



BACK TO CONTENTS

"Between security challenges and potential benefits, decision-makers **must strike a careful balance in their pursuit of GenAI."**

## Despite the High Prioritization of GenAI, Organizations Face a Diverse Set of Challenges

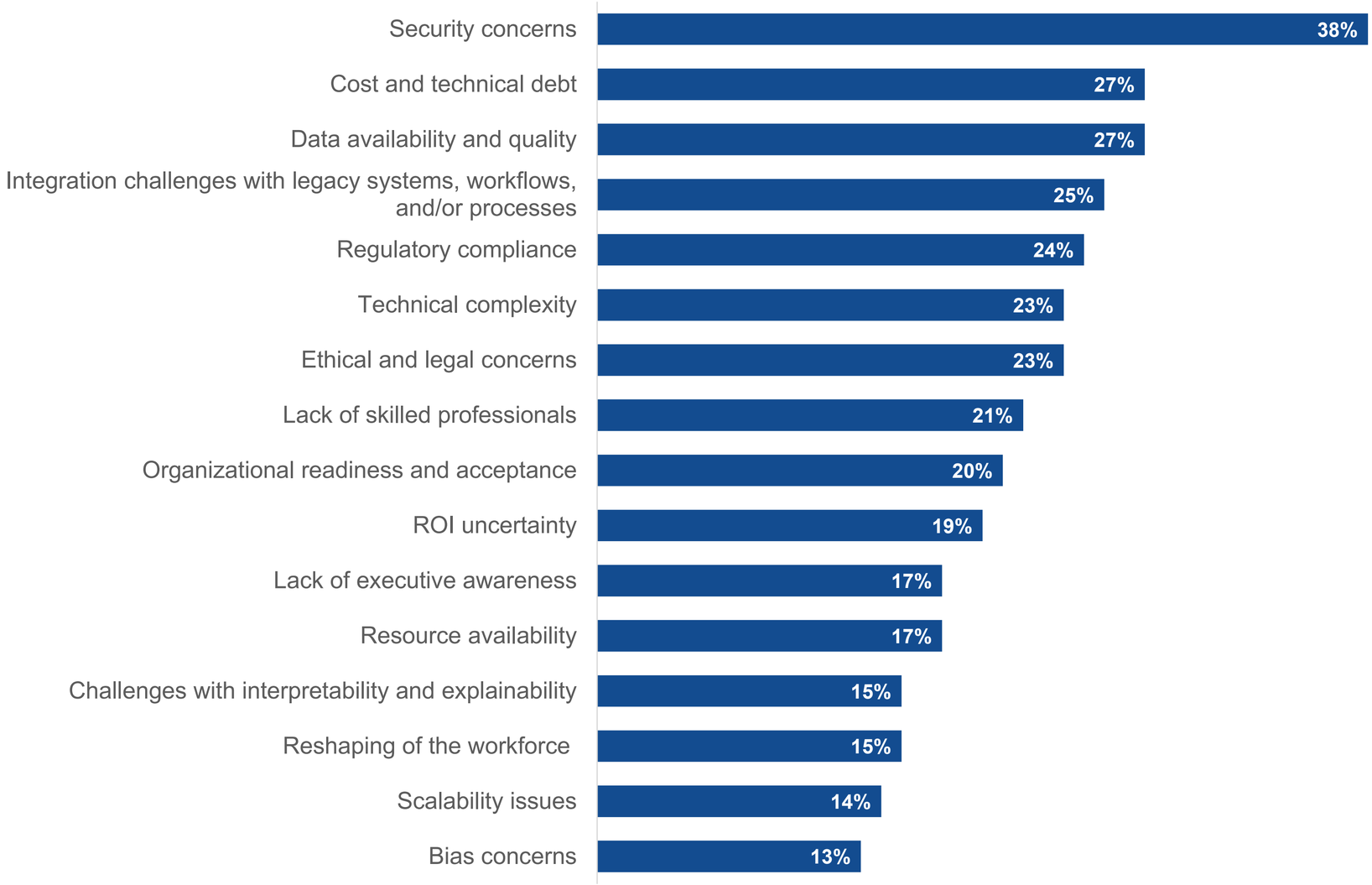
While organizations on the sidelines grapple with challenges related to cost, data quality, and daunting infrastructure integration/modernization requirements, security remains the top roadblock.

GenAI security remains a hot topic as organizations seek to overcome challenges such as data privacy, confidentiality, potential misuse/abuse, data leaks, unauthorized access, and more.

But while some view security concerns as a challenge, there is a growing trend of folks embracing the potential of GenAI to enhance certain security measures and tasks, such as providing advanced threat detection and enhancing automated response systems.

Between security challenges and potential benefits, decision-makers must strike a careful balance in their pursuit of GenAI.

**Challenges Preventing the Adoption of GenAI**






BACK TO CONTENTS

"On average, respondents believe ~44.3% of their **employees are currently using GenAI in some capacity."**

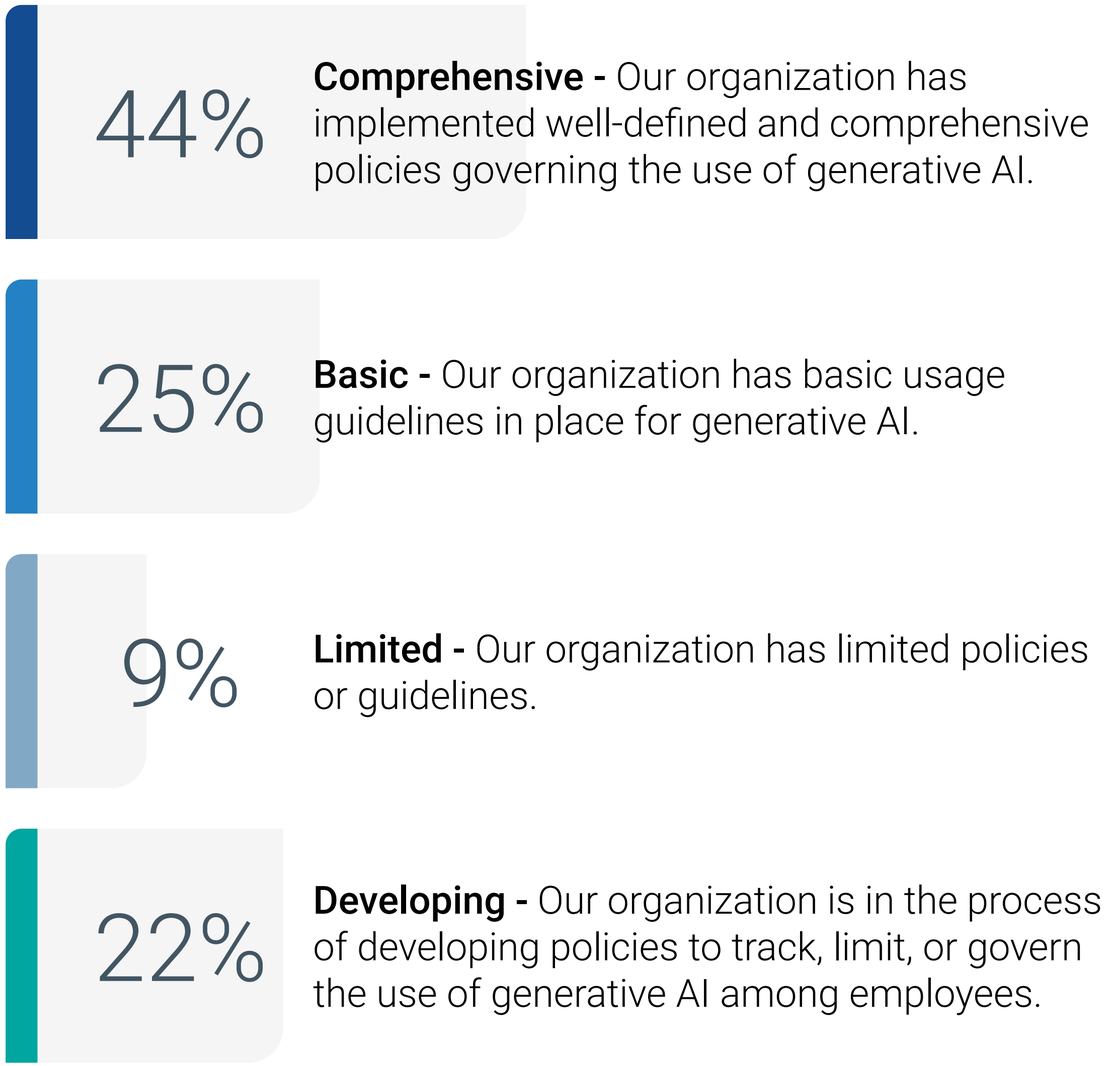
## The Imbalance Between Usage and Company Governance

As organizations ramp up internal GenAI initiatives, they often find employees are already leveraging free online tools, such as ChatGPT.

Less than half of organizations have implemented comprehensive policies to define the use or governance of openly available GenAI technologies. There is still a lot of work for organizations to do to ensure they properly monitor what employees are doing with generative AI.

In addition, 61% of organizations agree that few people know how to use GenAI properly or how to fact-check the result. With organizations slow to put usage guidance or policies in place, it's not surprising that 57% of organizations find it challenging to know who is leveraging GenAI in their workflows.

**The Extent of GenAI Usage Guidance or Policy Implementation**





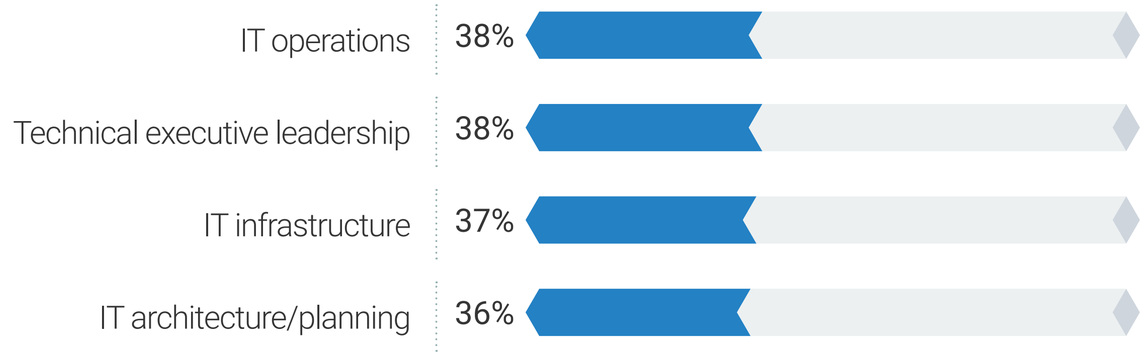

BACK TO CONTENTS

# The Enterprise Stakeholder Landscape for GenAI

# Generative AI Decision-makers and Budget Holders

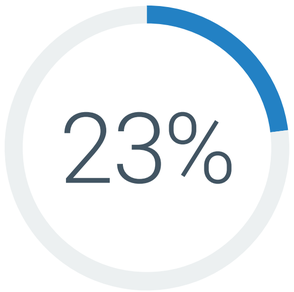
While influence in GenAI decisions spans a wide variety of roles and departments, IT leadership is rapidly rising to the top. Many IT organizations are already tasked with experimenting and evaluating the potential of emerging technologies, and GenAI is no different. IT has traditionally owned most, if not all, of enterprise technology spend. With GenAI being so accessible to consumers, a new challenge has presented itself. Business leaders are not being empowered to use their own budgets to consume GenAI services. And this shift has left several stakeholders within the business asking, "Who else do I need to involve before I make a decision?" and "Who is directing budget to this initiative?"

## Stakeholders Influencing GenAI Purchase Decisions

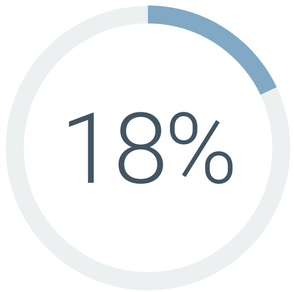



## Primary Budget Holders


IT leadership

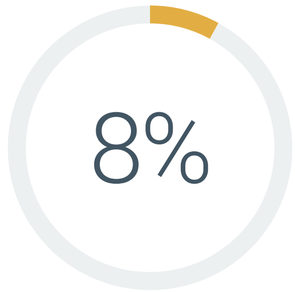



Technical executive leadership




Data team leadership


Line of business leader




Non-technical executive leadership



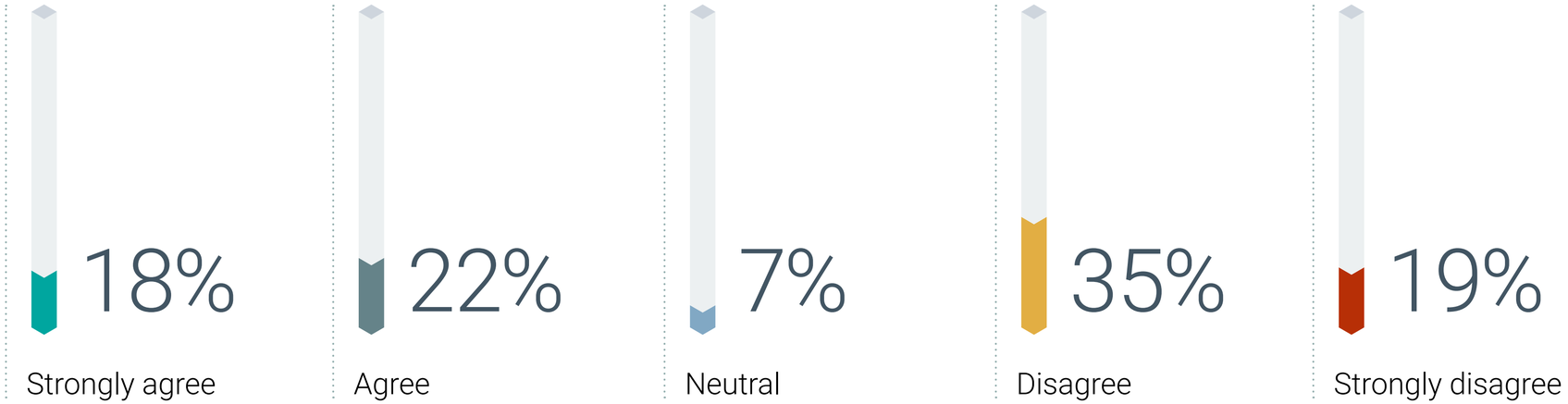
BACK TO CONTENTS

# The Growing Need to Educate the Wider Stakeholder Ecosystem Involved in GenAI Projects

Ensuring that decision-makers and relevant stakeholders are adequately informed about GenAI projects is crucial for successful planning and execution. The issue today is that 40% of organizations believe their involved parties are not well informed, painting an ugly picture for the state of the business.

With an increasing number of stakeholders involved in decision-making for GenAI initiatives, there is growing need for education across the business. Education extends to collaboration, with most organizations having significant room to improve in this area.

## The Stakeholder Disconnect




**40%**
of organizations believe their involved parties are not well informed, **painting an ugly picture for the state of the business.**



BACK TO CONTENTS

"88% of organizations said they **have started cross-functional collaboration among IT, data science, and business teams** to better utilize GenAI within their business."

## There Is Always Room to Improve Collaboration

Over a third (39%) of organizations indicated they have excellent collaboration across teams responsible for selecting, implementing, and managing GenAI solutions. This number increases to 45% among C-level executives, compared to 36% of other managers and practitioners.

This need to improve collaboration is a big reason why 88% of organizations said they have started cross-functional collaboration among IT, data science, and business teams to better utilize GenAI within their business.

**Rating the Level of Collaboration Among Decision-makers, Implementers, and Ops Teams**


**Excellent -** Teams collaborate seamlessly and frequently share information and resources.


**Good -** Teams generally collaborate well, but there is room for improvement in sharing information and resources.


**Fair -** Collaboration between teams is average, with occasional sharing of information and resources.


**Poor -** Teams struggle to collaborate effectively, resulting in limited sharing of information and resources.



BACK TO CONTENTS

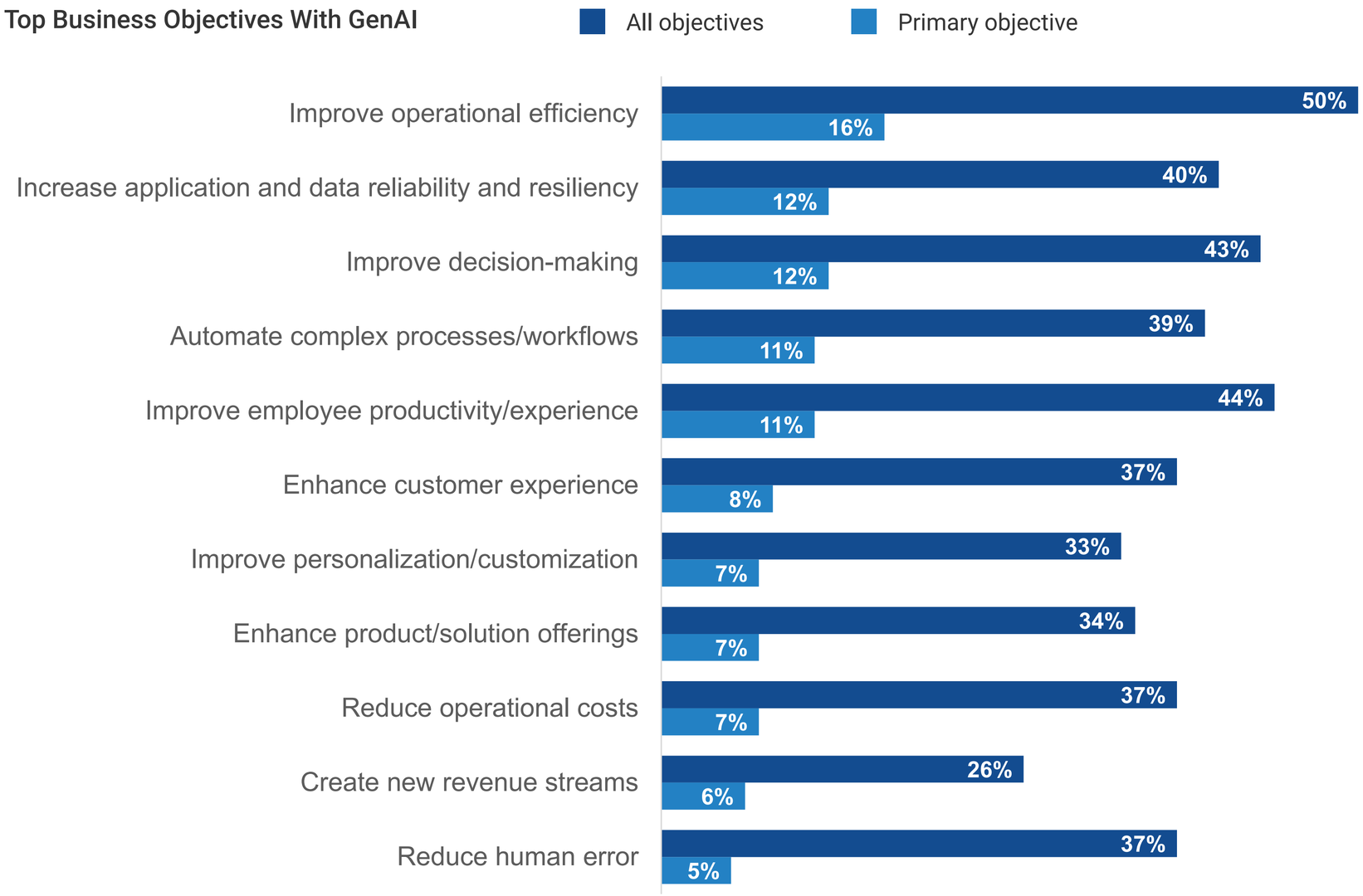
# Enterprise Objectives, Use Cases, and Benefits Driving Adoption

# Potential Improvements in Operational Efficiency Are Most Often Driving GenAI Initiatives

As organizations pursue GenAI, they see several areas of the business where they believe it could help; the most common is in improving operational efficiency.

With so many approaches and use cases, objectives will vary by size of organization, level of existing expertise, industry, and persona. The ability of an organization to prioritize and focus on the right use cases will be paramount to finding early success with generative AI.

Interestingly, when looking at the research based on persona, IT stakeholders—who are often overburdened with too many requests and not enough time—are particularly interested in improving productivity, reducing human error, and enhancing the customer experience.





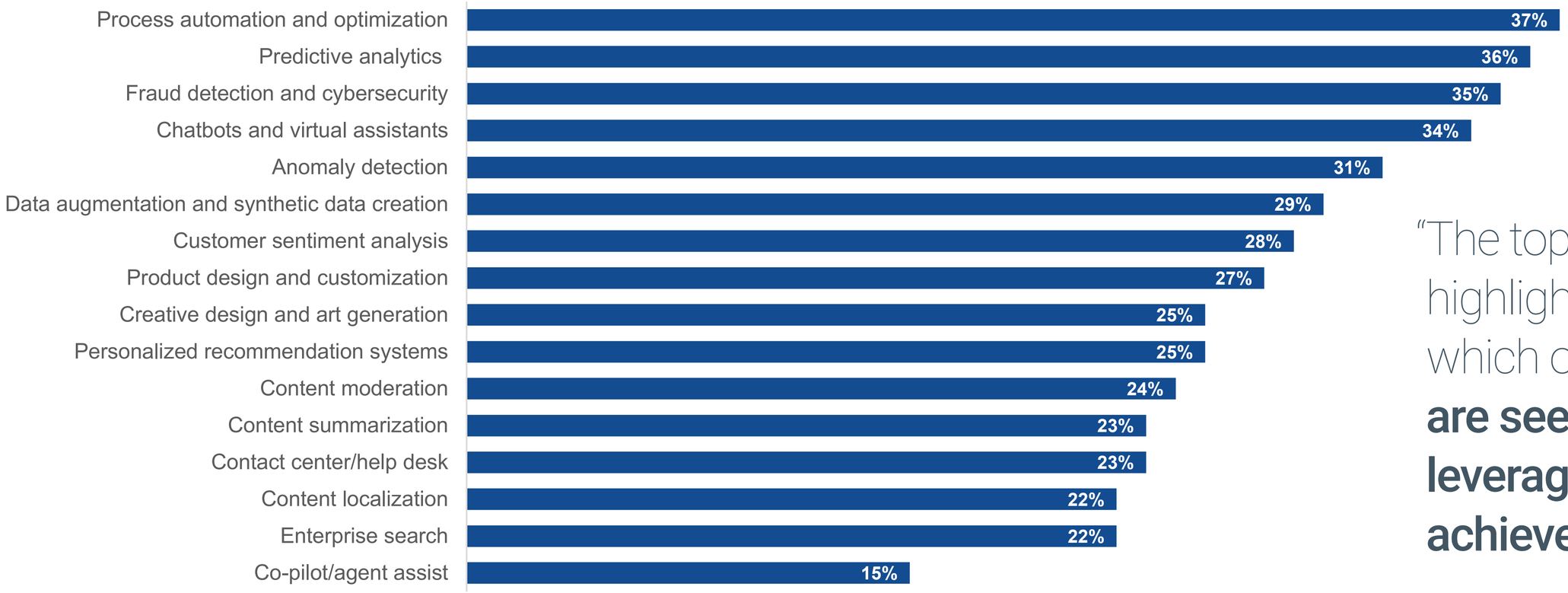

BACK TO CONTENTS

# Diverse Set of Use Cases Drives Organizational Investments

The top use cases highlight the ways in which organizations are seeking to leverage GenAI to achieve objectives, with leading use cases focused on the automation and optimization of process, analytics, and cybersecurity.

An additional level of complexity gets added to the use case landscape based on key industries. For instance, in healthcare, multimodal GenAI can be utilized to analyze medical imaging and assist in diagnosing diseases with high accuracy. In finance, GenAI can enhance fraud detection by identifying unusual patterns and predicting potential fraud. Given the relative infancy of this space, the number of potential use cases is expected to increase significantly in the coming years, especially as organizations seek to build greater levels of trust and reliability in the accuracy of GenAI responses.

**Potential Enterprise Use Cases Driving Investment in Generative AI Solutions (Multiple Responses Accepted)**




"The top use cases highlight the ways in which organizations **are seeking to leverage GenAI to achieve objectives**."



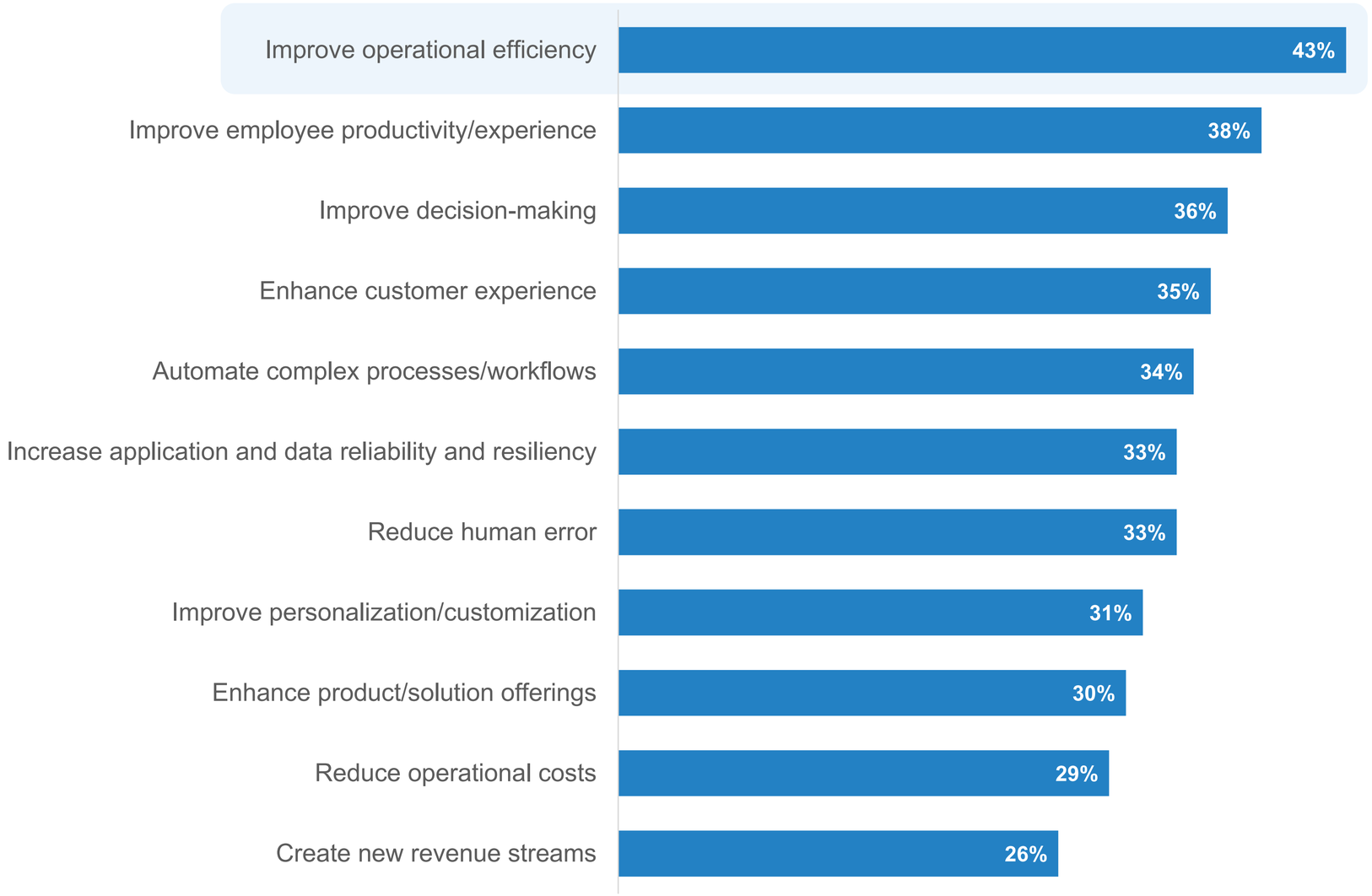
BACK TO CONTENTS

# Early GenAI Initiatives Are Delivering Results

Many organizations have already started using GenAI, and they've been able to see value rather quickly. And these realized benefits are often interconnected, as they collectively enhance operational efficiency, improve accuracy, and drive innovation across many domains.

Improving operational efficiency is cited as the top benefit realized from GenAI, mainly due to its substantial impact on productivity and cost savings. By automating routine tasks and optimizing complex workflows, GenAI is enabling organizations to reduce manual effort and minimize errors.

## Realized Benefits of GenAI






BACK TO CONTENTS

# The Success of GenAI Initiatives Is Measured From Multiple Perspectives

Value means something different depending on the business, group, industry, stakeholder, and customer. It's important to understand that organizations have different ways of defining value and measuring success.

## Measuring GenAI Success

43%

Qualitative impact analysis

43%

Overall accuracy of generative AI responses

40%

User/process quantitative benefits

38%

Cost-savings/return on investment

36%

Adoption rates and the integration of AI technologies into existing workflows and processes

33%

Industry and/or domain-specific key performance indicators (KPIs)

31%

Measuring results against industry standard benchmarks

31%

Risk mitigation and compliance

31%

Employee skill development

29%

Feedback loops and iterative improvement to assess and refine outcomes



BACK TO CONTENTS

## Conclusion

The research presented in this report underscores the transformative power of GenAI for enterprises. From streamlining operations to fostering creative innovation, GenAI holds immense potential to reshape entire industries. Key trends, such as advancements in natural language processing and the rise of multimodal models, pave the way for even broader applications. However, navigating this complex ecosystem requires careful consideration.

***Enterprise Infrastructure: The Unsung Hero*** - This report emphasizes the critical role of infrastructure in enabling successful GenAI adoption for enterprises. Our survey findings reveal that a robust, scalable, and secure infrastructure foundation is essential for maximizing GenAI performance and building trust within organizations.

***The Road Ahead*** - As GenAI continues to evolve, ongoing research and collaboration among stakeholders will remain crucial. By building a strong infrastructure foundation and carefully navigating data and model selection, enterprises can unlock the true potential of GenAI and secure a competitive advantage.

Overall, this report equips readers, particularly enterprise decision-makers, with the knowledge and insights necessary to navigate the GenAI landscape effectively and forge a path toward successful GenAI implementation.

## Message from the Sponsor

Hitachi Vantara is transforming the way data fuels innovation. A wholly owned subsidiary of Hitachi Ltd., Hitachi Vantara provides the data foundation the world's leading innovators rely on. Through data storage, infrastructure systems, cloud management and digital expertise, the company helps customers build the foundation for sustainable business growth. To learn more, visit www.hitachivantara.com.

LEARN MORE

Hitachi Vantara

## RESEARCH METHODOLOGY AND DEMOGRAPHICS

To gather data for this eBook, Enterprise Strategy Group conducted a comprehensive online survey of 800 IT leaders (50%), line-of-business leaders (24%), and data-centric practitioners/management (25%) involved in generative AI decision-making.

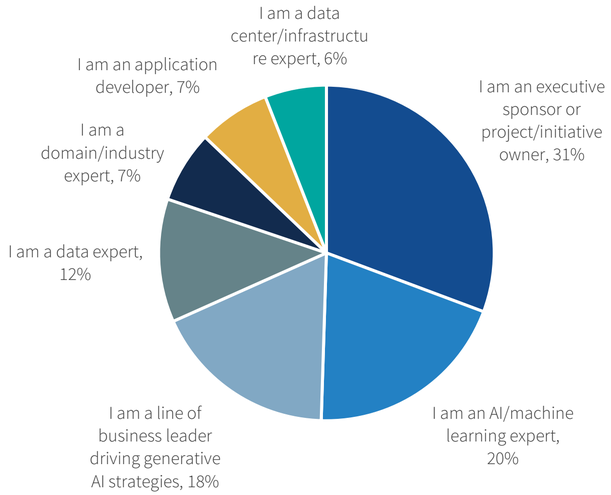
Organizations represented span all private- and public-sector verticals, including technology, telco/media, financial services, manufacturing, healthcare/life sciences, and retail/wholesale, among others. The research spanned both large midmarket (500 to 999 employees, 11%) and enterprise (1,000+ employees, 89%) organizations.

The research was global, including respondents based in North America (U.S. and Canada, 50%) and Western Europe (U.K., France, Germany, 50%). The survey was fielded between Jan. 17, 2024, and Feb. 16, 2024.

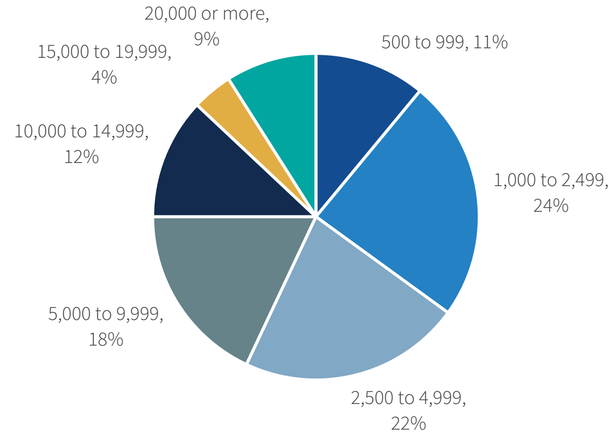
The margin of error at the 95% confidence level for sample sizes of n=800 are + or – 4 percentage points, respectively.

Note: Totals in figures and tables throughout this report may not add up to 100% due to rounding.

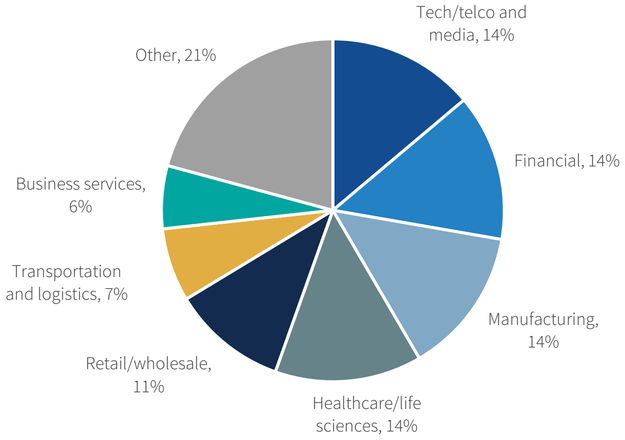
### Respondents by Generative AI Stakeholder Category



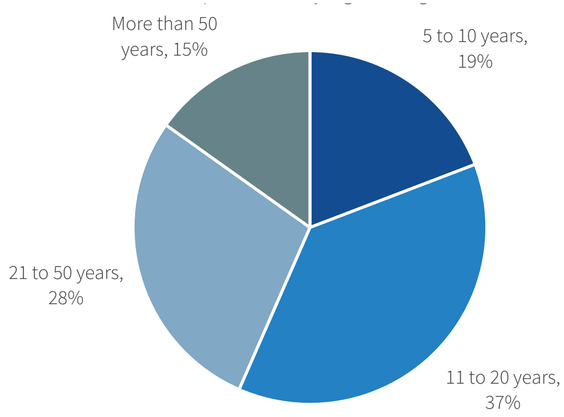

### Respondents by Number of Employees




### Respondents by Industry




### Respondents by Age of Organization






BACK TO CONTENTS

Enterprise Strategy Group™
by TechTarget

**Enterprise Strategy Group** is an integrated technology analysis, research, and strategy firm providing market intelligence, actionable insight, and go-to-market content services to the global technology community.