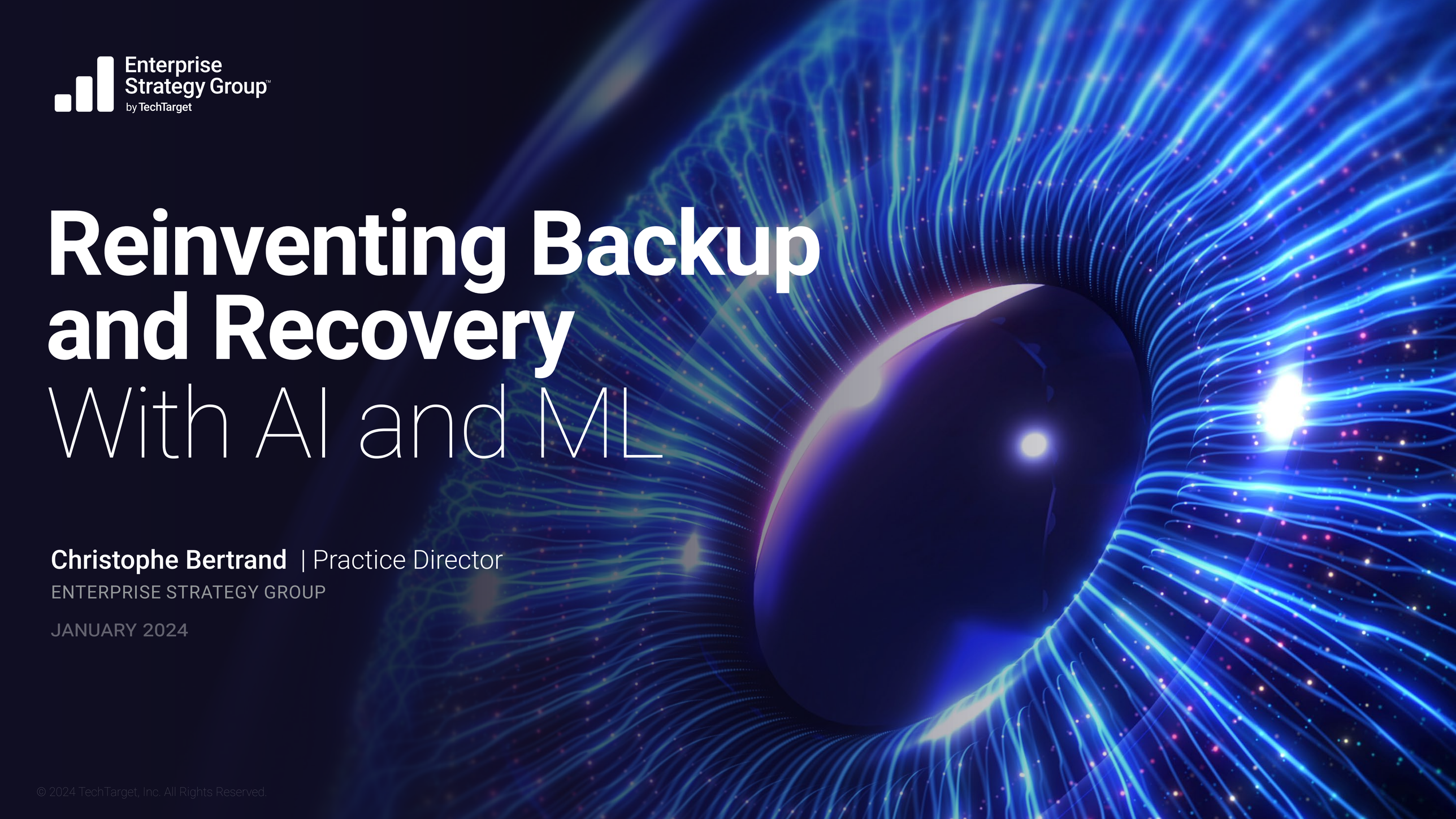


Enterprise Strategy Group™ by TechTarget

Reinventing Backup and Recovery With AI and ML

Christophe Bertrand | Practice Director
ENTERPRISE STRATEGY GROUP

JANUARY 2024

Research Objectives

Artificial intelligence (AI) is a topic frequently discussed in various contexts today, primarily regarding its impact on society and its application in specific scenarios. When it comes to backup and recovery, two questions are significant:

- **How much AI-generated data needs protection?**
  As more organizations integrate AI into their operations, the need to protect an increasing amount of data assets becomes a crucial concern. For example, it is important to assess which data sets are mission-critical and prioritize protection initiatives around them.

- **How will backup and recovery processes adapt to take advantage of these rapidly evolving technologies?**
  While some backup and recovery solutions already have AI and machine learning (ML) capabilities, further integration of AI/ML for autonomous data protection is receiving increased attention due to frequent data loss, often caused by criminal activities. Organizations are also carefully monitoring the inevitable impact of generative AI, which could drive significant expansion of data backup requirements.

To gain further insight into these trends, TechTarget's Enterprise Strategy Group surveyed 375 IT and data professionals familiar with and/or responsible for data protection (including backup and recovery) decisions and data science for their organization.

THIS STUDY SOUGHT TO:

Assess the state of the backup and recovery market in terms of AI and ML.

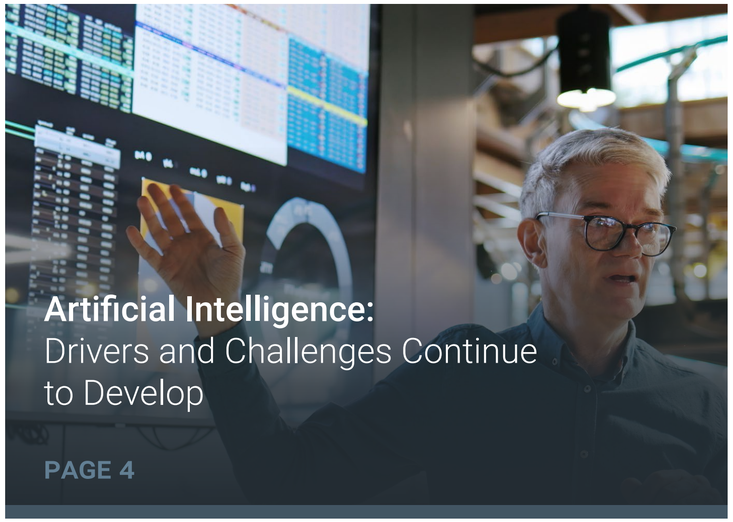
Uncover the impact of AI initiatives on backup and recovery processes and infrastructure, including the support of broader data initiatives.

Highlight current and future use cases for AI and ML in backup and recovery solutions.

Determine the stakeholders involved with the selection of backup and recovery platforms in the context of AI and ML, and understand the spending intentions and drivers for these technologies.

BACK TO CONTENTS

KEY FINDINGS

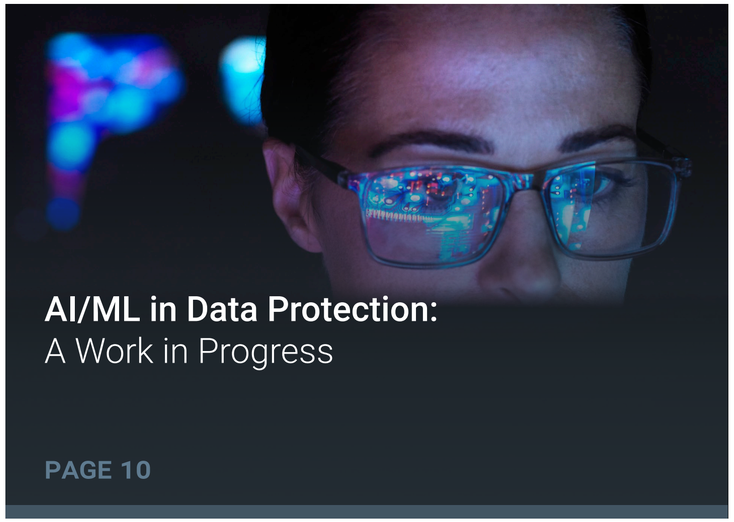


Artificial Intelligence:
Drivers and Challenges Continue to Develop

PAGE 4

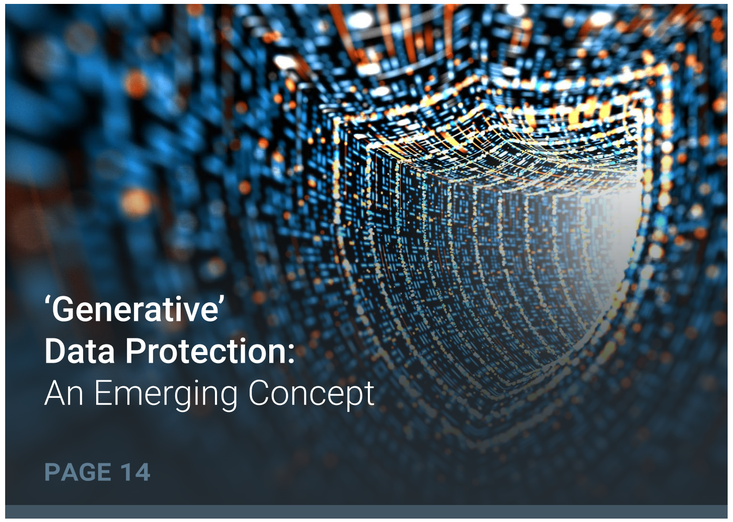
The AI Data Backup Gap:
A Look Inside

PAGE 7



AI/ML in Data Protection:
A Work in Progress

PAGE 10



‘Generative’ Data Protection:
An Emerging Concept

PAGE 14

Ransomware:
AI/ML to the Rescue

PAGE 16

Data Governance and AI:
It’s Complicated

PAGE 18

BACK TO CONTENTS

Artificial Intelligence: Drivers and Challenges Continue to Develop

"In time, organizations will focus more on **customer experience and engagement**, as this is only the beginning of new era."

AI Is Primarily Internally Focused

The primary business objectives for implementing AI are various, spanning a vast array of use cases and expected outcomes. Overall, while many organizations are seeking to leverage technology and data to expand their business success, customer and go-to-market considerations are not the top priority today. The prism is primarily dedicated to operational efficiency and profit enhancements as well as better leveraging and managing data as an asset. In time, organizations will focus more on customer experience and engagement, as this is only the beginning of new era.

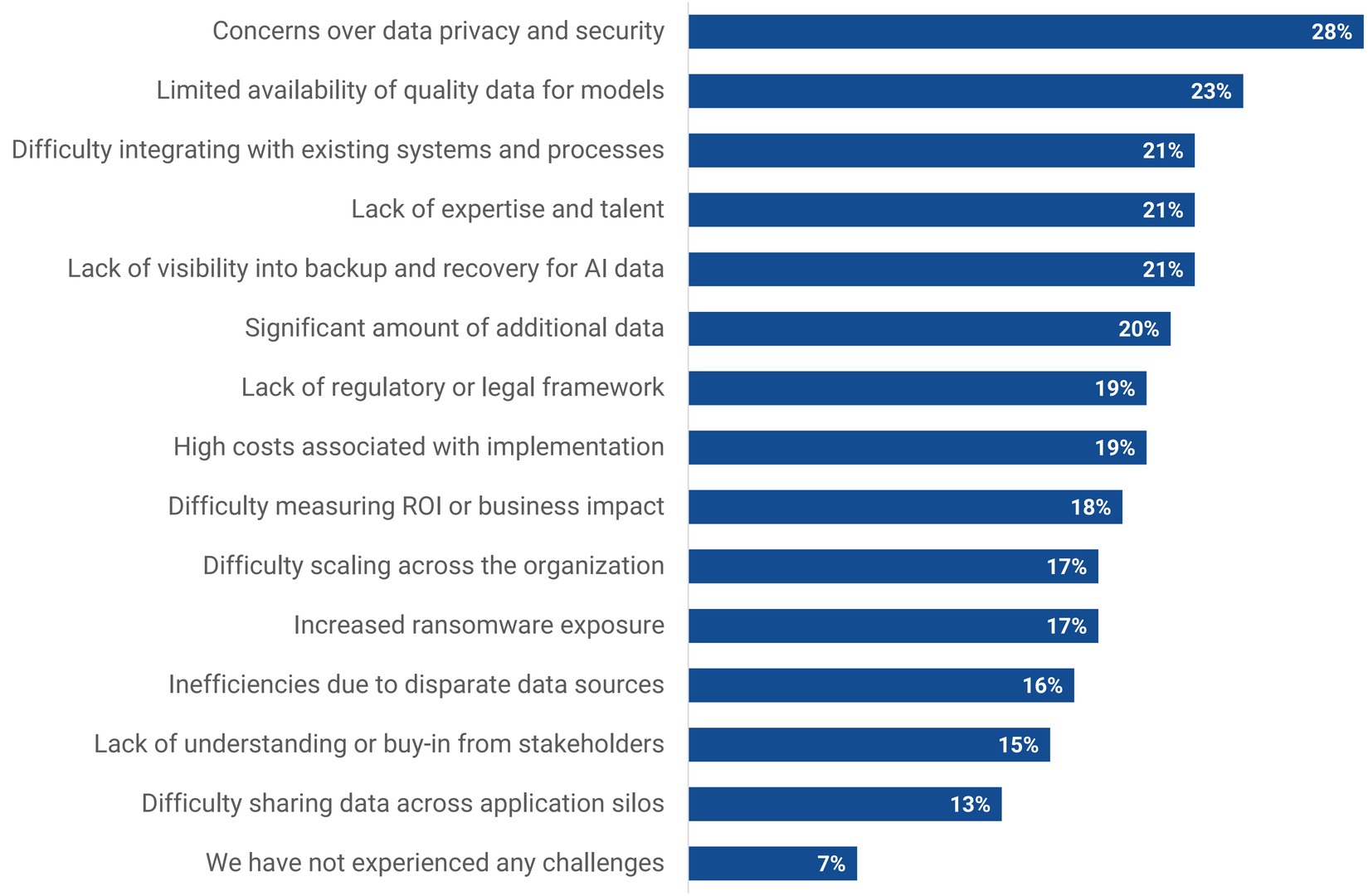
Business objectives for implementing AI.

BACK TO CONTENTS

Growing Pains Challenge Organizations on Their AI Journey

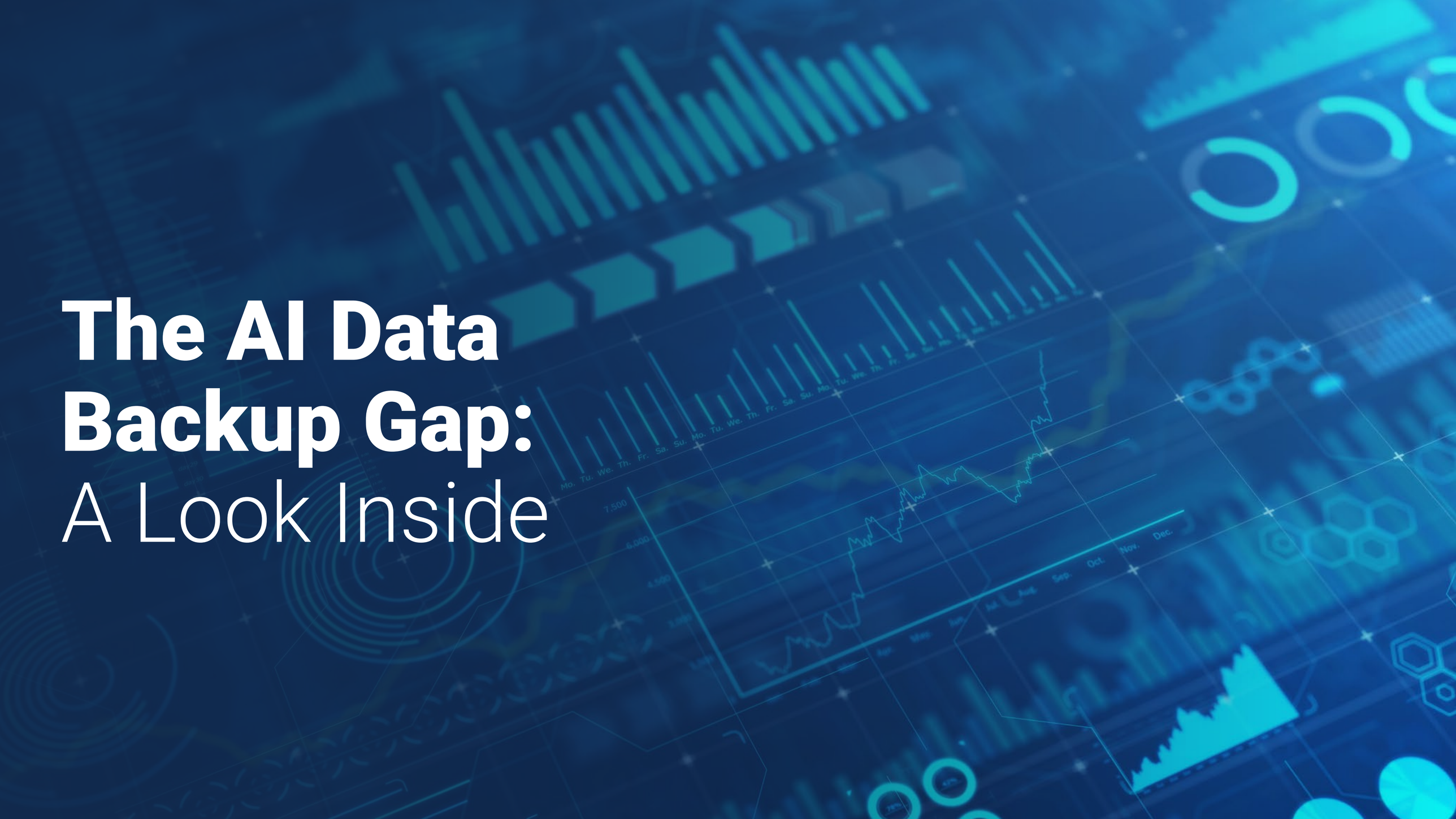
Challenges abound across business and technology as AI initiatives become mainstream. Broadly, these challenges comprise four primary areas: compliance, security, data management, and deployment. While these are different fields by nature, in many cases, the challenges are intertwined. At the top of the list are data privacy compliance and security, which are fundamental challenges hindering deployments, a problem likely compounded by skill sets shortages.

Challenges encountered when implementing AI.



BACK TO CONTENTS

The AI Data Backup Gap: A Look Inside



Don’t Lose AI Data… or Else

In the realm of traditional applications and workloads, losing data of any kind has never been a good thing for any organization. The same logic applies to AI, with the specter of data recreation costs at the top of the list of consequences. Backup vendors will be on the frontlines as organizations grapple with the consequences of lost AI data that was not backed up or is otherwise unrecoverable. This also shows that while AI may be in its early stages, organizations are not discounting backup and recovery. Historically, that is significant simply because newer workloads tend to be an afterthought from a backup standpoint.

Consequences of losing AI data that was not backed up or is otherwise unrecoverable.

High cost incurred for recreating the models

Reevalute our AI vendor(s)

Reengineer data backup plans and processes

Reevaluate our backup vendor(s)

Consult third-party advisors

Organizational changes

“Backup vendors will be on the frontlines as organizations **grapple with the consequences of lost AI data that was not backed up** or is otherwise unrecoverable.”

Christophe Bertrand,
Practice Director & Principal Analyst

BACK TO CONTENTS

The AI Data Backup Gap

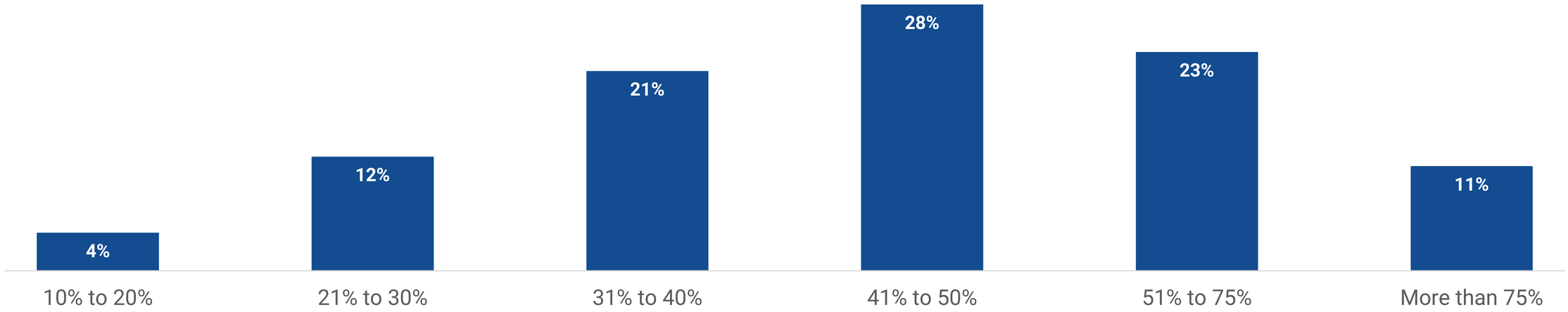
A large number of organizations are not doing enough to protect their AI-generated data. Shockingly, 65% admit to regularly backing up only up to 50% of their total volume of AI-generated data. This is a major concern given the nature of the workload and investment therein, especially regarding the protection of data assets, intellectual property, and the value built in data models. Furthermore, AI processes are rapidly becoming integral to production processes, if not the actual business itself. This means that the value of AI-generated data is on the rise, and so is the need to protect it.

Although many organizations have yet to take stock of the criticality of backup and recovery for AI-generated data, the picture should improve moving forward. As businesses continue to realize the importance of AI in their operations, they will also realize the importance of protecting the data generated by these systems. In the meantime, the current gap in backup and recovery for AI-generated data is a major concern, as it poses significant business risks and compliance exposures. Organizations that fail to protect their AI-generated data risk losing out on valuable insights, intellectual property, and competitive advantage. Therefore, it is imperative that businesses start taking steps to safeguard their AI-generated data before it's too late.

"AI processes are rapidly becoming **integral to production processes**, if not the actual business itself."

65%
admit to regularly backing up **only up to 50% of their total volume** of AI-generated data.

Percentage of AI-generated data included in regular backups.



BACK TO CONTENTS

AI/ML in Data Protection: A Work in Progress

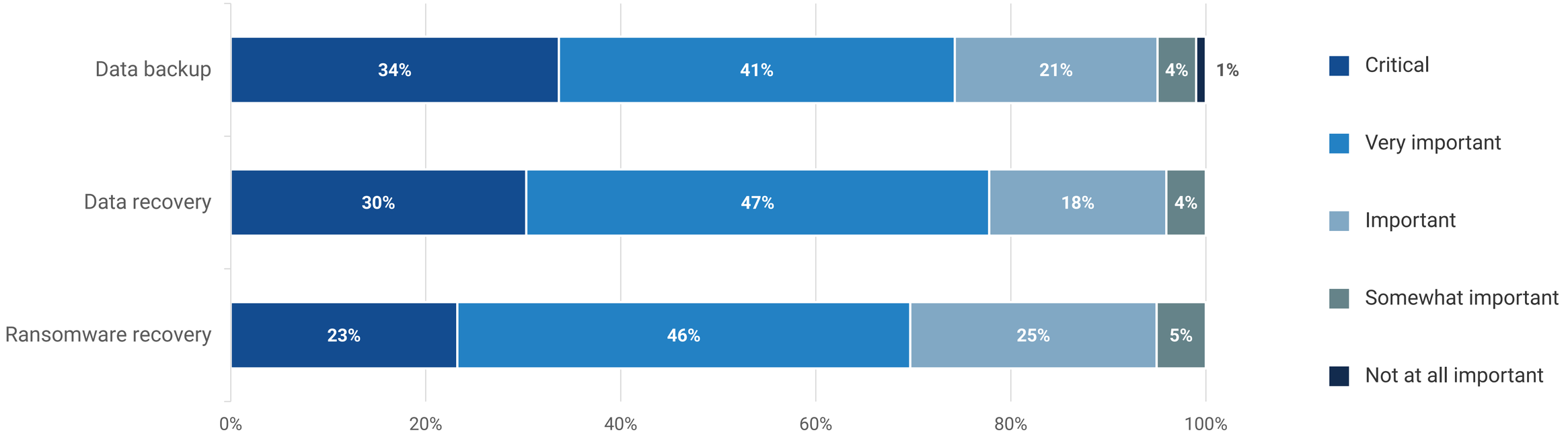
Over 90%

of organizations report that **AI and ML are important.**

AI/ML Seen as Essential to Modern Data Protection Processes

Over 90% of organizations report that AI and ML are important, if not very important or critical, to their data backup, data recovery, and ransomware recovery efforts, foreshadowing significant progress in this space moving forward.

Future importance of AI/ML technology for data protection processes.

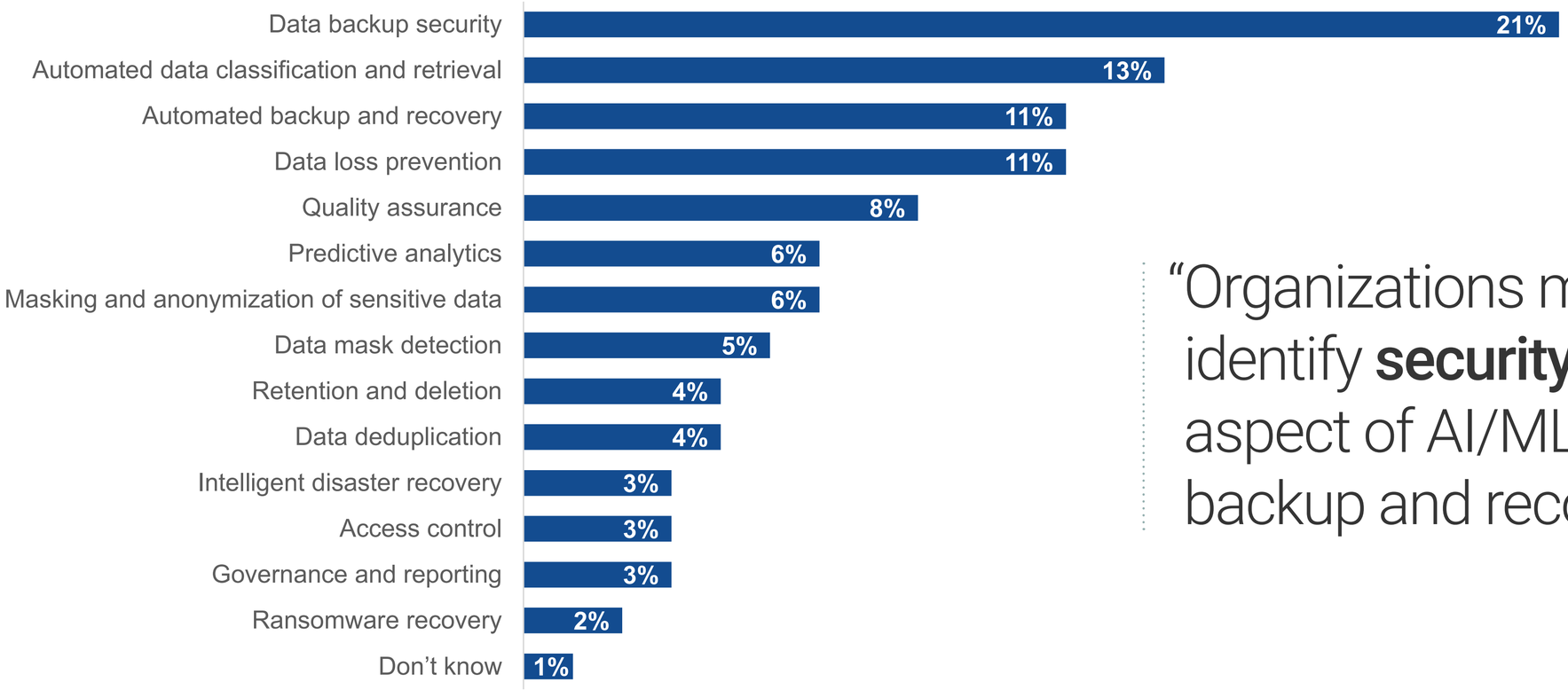


BACK TO CONTENTS

Enhanced Security Is the Big Prize From AI/ML Integration

Integrating or furthering the use of AI/ML in backup processes is expected to support multiple discrete yet data-centric use cases. It should be no surprise that organizations most commonly identify security as the top beneficial aspect of AI/ML integration into backup and recovery, followed by automation and data loss prevention. Also notable is the emergence of data management technology with automated data classification and retrieval, which allow for data reuse and data compliance for security posture management, for example.

Use cases for AI/ML in data backup and recovery expected to have the most positive impact.



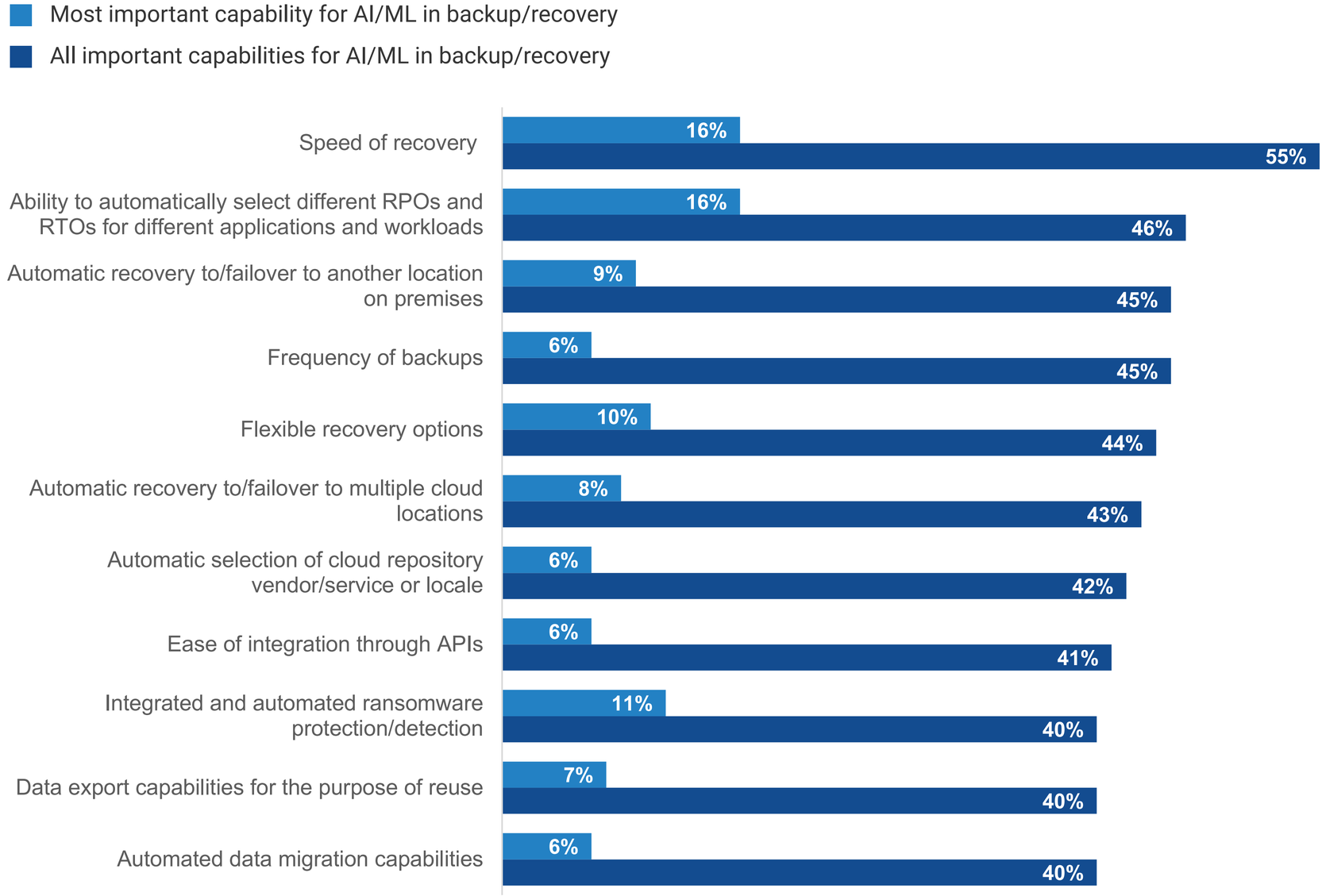
“Organizations most commonly identify **security** as the top beneficial aspect of AI/ML integration into backup and recovery.”

BACK TO CONTENTS

The ‘RFP’ List for AI/ML in Backup and Recovery

It’s no secret that AI and ML are becoming an increasingly important part of backup and recovery solutions for end users. While many vendors have already incorporated these technologies into their offerings, this space needs more innovation. One of the key requirements for backup and recovery solutions is the ability to provide both “traditional” and augmented capabilities. This means that organizations need to be able to recover data quickly and easily, but they also want advanced features that can improve the overall efficiency and effectiveness of their backup and recovery processes. In addition to these capabilities, organizations are looking for solutions that can help them prepare for the growing threat of ransomware attacks. With the rise of these attacks in recent years, ransomware preparedness integrations have become a critical requirement for backup and recovery solutions.

Important and *most* important capabilities for AI/ML in data protection.



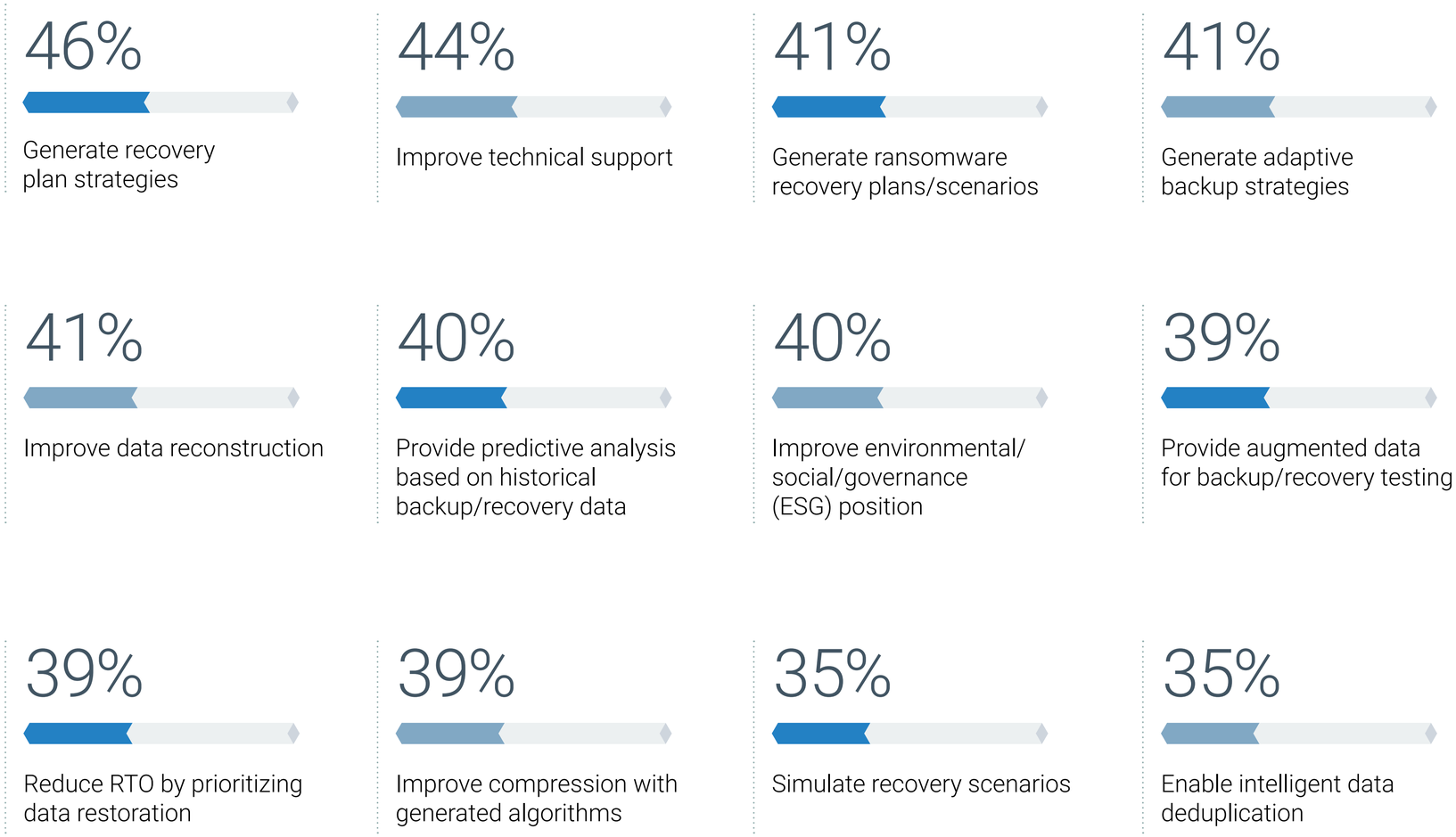
BACK TO CONTENTS

‘Generative’ Data Protection: An Emerging Concept

Expectations Are High for Generative AI Use in Backup and Recovery

Leveraging AI-based natural language interfaces to solve complex issues or troubleshoot expert processes is becoming central to many organizations' support strategies. This is true in the space of backup and recovery as well, with a twist: While support is key, generating the actual recovery plan strategy—a very significant added value—is the most cited way generative AI is expected to improve backup and recovery. Of note are also the testing aspects associated with the battle against ransomware. This highly complex and highly valued area must be addressed. All these expected benefits provide another way to gauge the anticipated generative AI capabilities vendors should build into their solutions.

Expected ways generative AI will improve data backup and recovery processes.



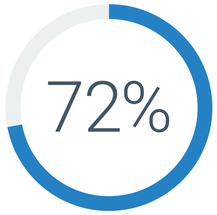
BACK TO CONTENTS

Ransomware: AI/ML to the Rescue

The Panacea for Recovery SLAs?

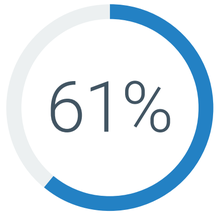
Ransomware is an existential issue for most organizations, well beyond just IT or security teams. With the business at stake, organizations have been deploying new capabilities to fend off and recover from attacks, and leveraging integration and automation is key. Among organizations that feel AI/ML will improve their ability to recover from ransomware attacks, more than two-thirds feel its automation will improve their overall cybersecurity-recovery RPO and RTO, among others benefits.

Future importance of AI/ML technology for data protection processes.



Improve overall cybersecurity-recovery RPO and RTO

Improve network traffic analysis

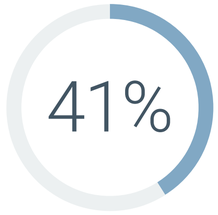


Automate intelligent data recovery

Improve compliance posture

Improve user behavior analysis

Enable AI-powered sandboxing



Lower cybersecurity insurance costs

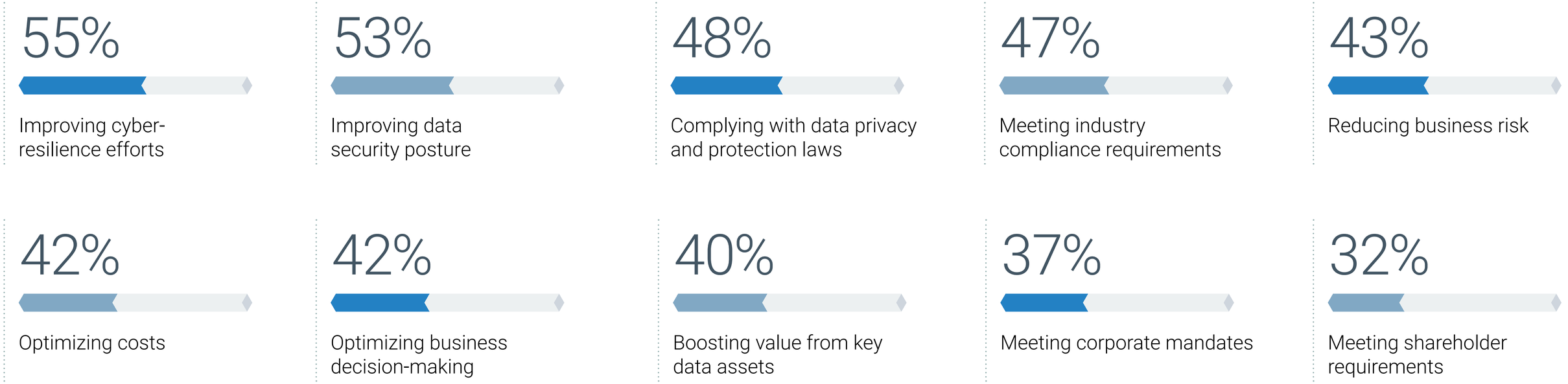
BACK TO CONTENTS

Data Governance and AI: It’s Complicated

Responsible, Governed AI Means Security and Cyber Resilience First

As organizations deploy more AI-driven processes and applications, data governance programs are moving into a new critical dimension. A variety of drivers are concurrently at play, with cyber resilience as the top motivation for a majority of organizations, reflecting a sign of the times. Examining the entire picture, most top drivers are “defensive” in nature and focus on meeting requirements and mandates as opposed to supporting growth initiatives. Watching how this ranking evolves over time as initiatives mature will serve as a compelling and crucial activity.

Business drivers that most influence data governance programs and responsible AI use.



BACK TO CONTENTS

Data Governance Initiatives Spark Growing Pains

Multiple types of challenges are in play simultaneously with data governance because there's just too much data. In what demonstrates the early stages of adoption, the perfect storm is forming: Lack of strategy, too much data, and too many applications, compounded by siloed management. Also notable is that the burden of compliance and its inherent complexity can further hinder the deployment of data governance initiatives. In time, these scale-related issues should abate as end users become more adept at managing their governance programs at scale.

Challenges faced when implementing and managing data governance initiatives and responsible AI use.

BACK TO CONTENTS

COHESITY

ABOUT

Cohesity is a leader in AI-powered data security and management. Aided by an extensive ecosystem of partners, Cohesity makes it easy to secure, protect, manage, and get value from data — across the data center, edge, and cloud. Cohesity helps organizations defend against cybersecurity threats with comprehensive data security and management capabilities, including immutable backup snapshots, AI-based threat detection, monitoring for malicious behavior, and rapid recovery at scale. Cohesity solutions can be delivered as a service, self-managed, or provided by a Cohesity-powered partner. Cohesity is headquartered in San Jose, CA and is trusted by the world's largest enterprises, including six of the Fortune 10 and 42 of the Fortune 100.

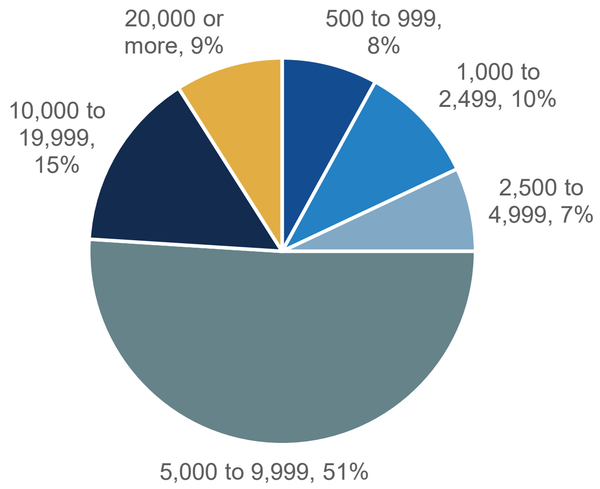
LEARN MORE

RESEARCH METHODOLOGY AND DEMOGRAPHICS

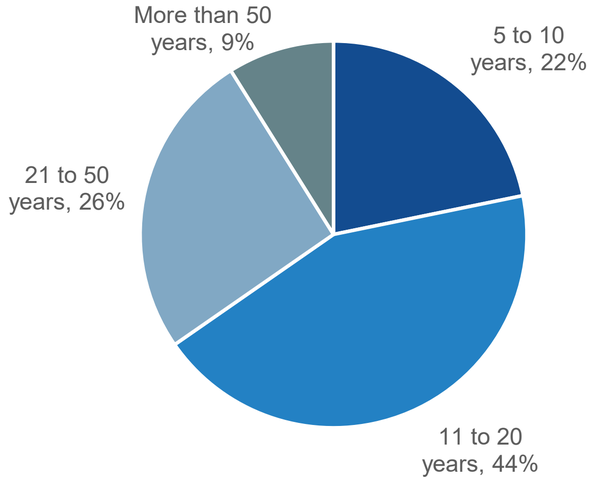
To gather data for this report, Enterprise Strategy Group conducted a comprehensive online survey of IT and data professionals from private- and public-sector organizations in North America (United States and Canada) between October 27, 2023 and November 17, 2023. To qualify for this survey, respondents were required to be familiar with and/or responsible for data protection (including backup and recovery) decisions and data science for their organization. All respondents were provided an incentive to complete the survey in the form of cash awards and/or cash equivalents.

After filtering out unqualified respondents, removing duplicate responses, and screening the remaining completed responses (on a number of criteria) for data integrity, we were left with a final total sample of 375 IT and data professionals.

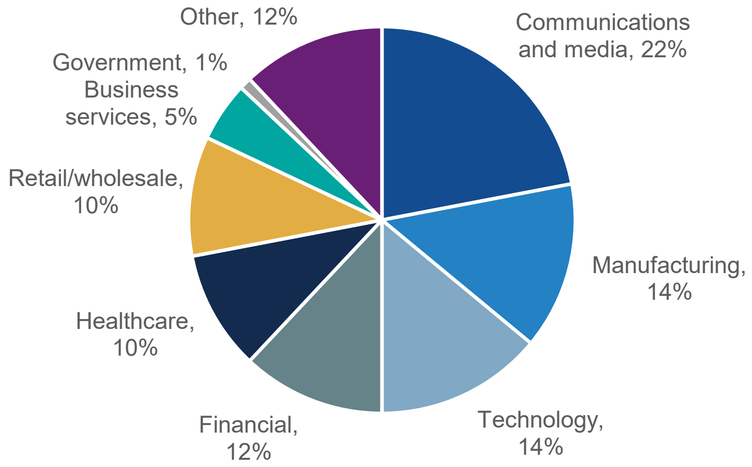
Respondents by Number of Employees



Respondents by Company Age



Respondents by Industry



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.