Enterprise Strategy Group | Getting to the bigger truth.™

RESEARCH HIGHLIGHTS

The Transformational Rise of Active Archives

Christophe Bertrand, Senior Analyst, Intelligent Data Management

AUGUST 2021

KEY FINDINGS

CLICK TO FOLLOW

Research Objectives

The digital economy is fueled by the seemingly unstoppable creation and insatiable consumption of data. To optimize costs, streamline management, and meet governance and regulatory requirements, organizations have traditionally archived non-production data on lower-cost storage tiers for years with little intention or need to reuse them. Things are evolving: The emergence and adoption of advanced digital transformation initiatives in recent years have placed data at the heart of the business and changed the nature of how archives can and should be leveraged. Rather than passive, "locked away" data sets, the past few years have seen the emergence of new requirements and solutions to make archives more "active" and leverageable to unlock business value.

In order to gain insight into these trends, ESG surveyed 150 IT professionals at organizations in North America (US and Canada) personally familiar with and/or responsible for data protection technology decisions, specifically data archiving and long-term retention strategies, for their organization. This research aimed to understand data archiving challenges, plans, strategies, and trends with a specific focus on active archive topologies whether on-premises or in the cloud.

THIS STUDY SOUGHT TO:

Assess the general state of data archiving environments and strategies, including benefits and challenges.

Highlight key trends and requirements with a focus on the impact of public cloud services on archiving.

Understand how IT organizations utilize active archive as part of their overall data archive strategy today.

Measure archive data growth and trends.

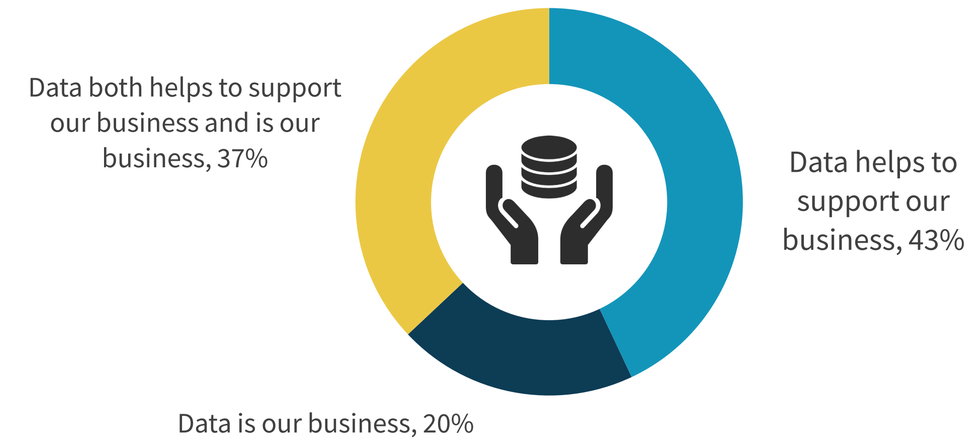
Back to Contents

Data Growth Is Not Slowing Down as More Businesses Become Data-reliant

Data-centric Future Ahead: Data Is, or Will Be, the Business

The acceleration of digital transformation has placed data at the heart of most organizations, private or public. In today's digital economy, most organizations leverage data as a product, a byproduct, or a source of enhancement of their offering. As a matter of fact, one in five report that data is purely their business. In today's economy, even organizations that traditionally produce tangible goods typically supplement these with online or digital offerings, subscriptions, etc. Indeed, among these organizations stating data helps to support their business, nearly two-thirds (62%) expect to develop data-centric products and or services in the next 24 months. Data is therefore truly becoming a business asset that can return value to the organizations that leverage it.

Organization perspectives on data importance.



62% of these organizations expect to develop **new data-centric products and services** in the next 24 months.

The mean total volume of data stored:

8.7 PB

The mean volume of archive data:

4 PB

Average length of time that archive information is retained:

10 years

Percentage of total archives that are active archives:

46%

Maximum length of time that archive information is retained:

15 years

Active archive annual mean growth rate:

36%

Back to Contents

General Data Archive Profile Guides Active Archive Strategy

The Three Dimensions of the Active Archive Profile

Long gone are the days of stale archives where data would be stored never (or very infrequently) to be used again—unless a governance or compliance issue required a recovery using a Ph.D. in archiving. Today, archives are accessed very frequently to the tune of multiple times per day, with 44% of respondents reporting daily access and 79% leveraging their archive on a weekly basis. This is amplified by another set of data points highlighting that 68% of IT professionals want to recover or retrieve data from their archive within minutes. These access characteristics are fundamental requirements to inform solution designs. The underlying technology to support these requirements must deliver from a performance and usability standpoint. Also, with half of the retrieved archive data being a year or less old, storage capacity should be planned accordingly. This dimension also influences archive solution designs.

In summary, archive access frequency, age of retrieved data, and archive access time are three key dimensions that can be used to identify profiles of users (and very likely use cases).

ACCESS FREQUENCY

Typical frequency with which information is retrieved or accessed from archived media/systems.

Daily
44%

Weekly
35%

AGE OF DATA

Average age of the information that is retrieved or accessed from archived media/systems.

A year or less
50%

TIME IT TAKES

Typical requirement among users in terms of archived information retrieval time.

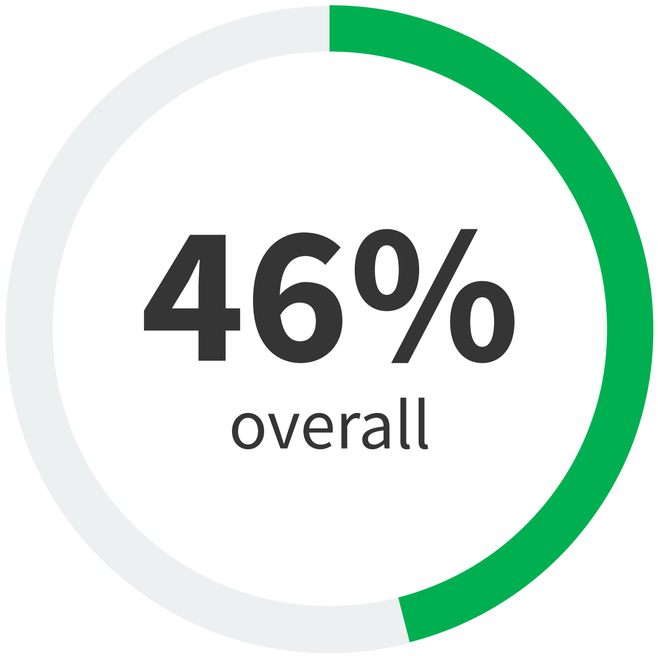
Within minutes
60%

Back to Contents

Organizations with Certain ‘Archive Profiles’ Are More Likely to Leverage Active Archive More Extensively

The average organization reports that about half (46%) of their total archive is considered active archive, meaning a tiered storage topology/solution that gives other IT systems or human end-users access to data through a common, unified file system that automatically retrieves and places that data on the appropriate storage tier. Based on the three dimensions of access frequency, access time, and age of retrieved data, we can determine certain organization profiles that are more likely to leverage active archive solutions more extensively. Specifically, organizations that store more than 10 petabytes of total archive data, retrieve archive data on a daily basis, typically retrieve data that is less than six months old, and want access to this data within seconds are likelier to have a higher percentage of active archive data.

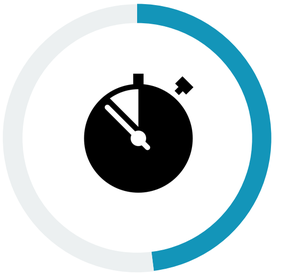
Percentage of total archive data that is active archive.



50%
for organizations with 10 PB or more of total archive data.

52%
for organizations that retrieve archive data on a daily basis.

52%
for organizations whose retrieved archive data is less than 6 months old.



48%
for organizations whose users need access to archive data within seconds.

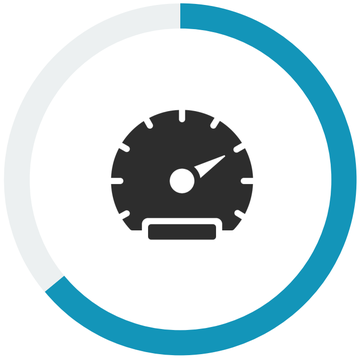
Back to Contents

Active Archives Are Significantly Improving Retrieval Times

Compared to legacy topologies, the adoption of active archive solutions has been met with significantly improved retrieval times for 64% of respondent organizations. Not surprisingly, this key benefit is directly correlated to spending intentions.

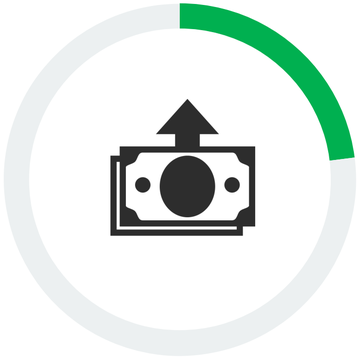
While overall nearly a quarter (23%) of organizations expect to significantly increase active archive spending over the next 24 months, among organizations that have seen archive data retrieval times significantly improve as a result of active archive, this jumps to 33%. Conversely, organizations that have seen little to no improvement expect much more conservative spending approaches when it comes to active archive.

Impact of active archive on data retrieval times compared to traditional archive solutions.



64%
say active archive has significantly improved data retrieval times.

Expected spending change for active archive strategy over the next 12 months.



23%
say it will Increase substantially.

“Organizations are bullish on active archive spending…especially among organizations that have seen improvements in data retrieval times.”

- Christophe Bertrand, ESG Senior Analyst

Percentage of organizations that expect to increase active archive spending substantially based on:

IMPROVEMENT IN DATA RETRIEVAL TIMES:

Significantly improved data retrieval times: **33%**

Slightly improved data retrieval times: **5%**

No change in data retrieval times: **7%**

Back to Contents

Active Archive Requirements and Benefits Commonly Involve Scale and Performance

Active Archive High Level RFP

IT professionals have identified the key features or characteristics they expect in a modern active archive. Not surprisingly, active archives require massive scale—the most important characteristic and often a core design parameter—combined with robust search capabilities at the price point that optimizes operational efficiencies. Active archives also distinguish themselves as solutions that support multiple types of storage technologies, including public cloud layers and on-premises deployment options.

Preferred attributes of active archive solutions.

65%
Has very high or unlimited storage scalability

53%
Provides contextual search capabilities

52%
Provides cost-efficiency

49%
Leverages public cloud services

47%
Provides a (front-end) user interface

46%
Must use a combination of storage technologies

39%
Must be available on-premises

35%
Provides a (front-end) performance layer

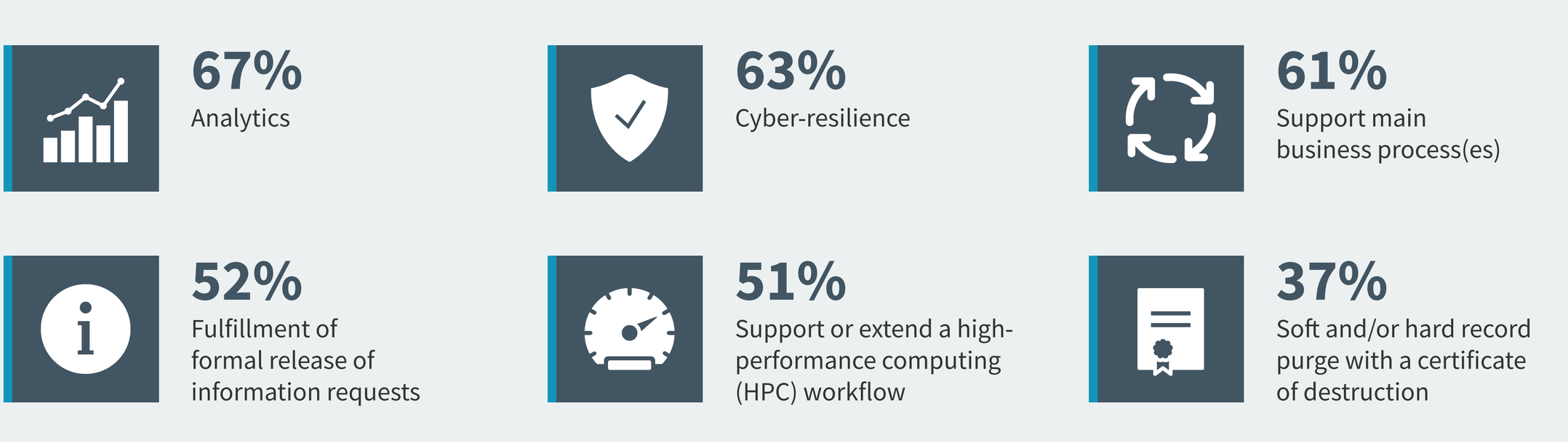
Back to Contents

Active Archive Use Cases

In the age of digital transformation and leveraging data assets, it is not surprising to see that analytics would top the list of active archive use cases. Also, active archives are a great complement to HPC type workflows (often very vertically integrated) in which a high-speed ingest and storage layer is typically complemented by other tiers of storage for further processing and cost optimization.

Also topping the list is cyber-resilience. Given the height and context of cyber risk with constant ransomware attacks, including to the data backups themselves, the ability to leverage archive data as an isolated recovery topology, for example, can offer some relief and participate in additional analytical processes.

Use cases supported by active archive solutions.



Back to Contents

Improved Data Access Is a Key Benefit of Active Archive

Organizations report having realized many key technical and business benefits with active archives. One particularly stands out: Improved data access, which is not only the most cited primary benefit but also the most common across all benefits.

While data access is very critical, it is closely followed by operational data efficiency-related advantages such as improved lifecycle or improved analytical process functions. In combination, these truly embody and confirm the data management value proposition of an "active" archive.

It is also important to highlight the "data is the business" advantages, with more efficiency in core business processes, improved ability to reuse data assets, improved abilities to innovate and compete, and the ability to not only create new products and offerings but also improve end-customer satisfaction. These are far-reaching and long-lasting benefits that can help an organization morph from "data reactive" to "data-centric."

For vendors in the space, the top 5 benefits should be highlighted in go-to-market messaging and proof-point development.

The top 5 benefits realized as a result of having an active archive strategy.

1. Improved data access

2. Improved data lifecycle

3. Operational efficiency in supporting core business

4. More efficient way to perform certain big data/analytics functions

5. Improved cyber-resilience

Back to Contents

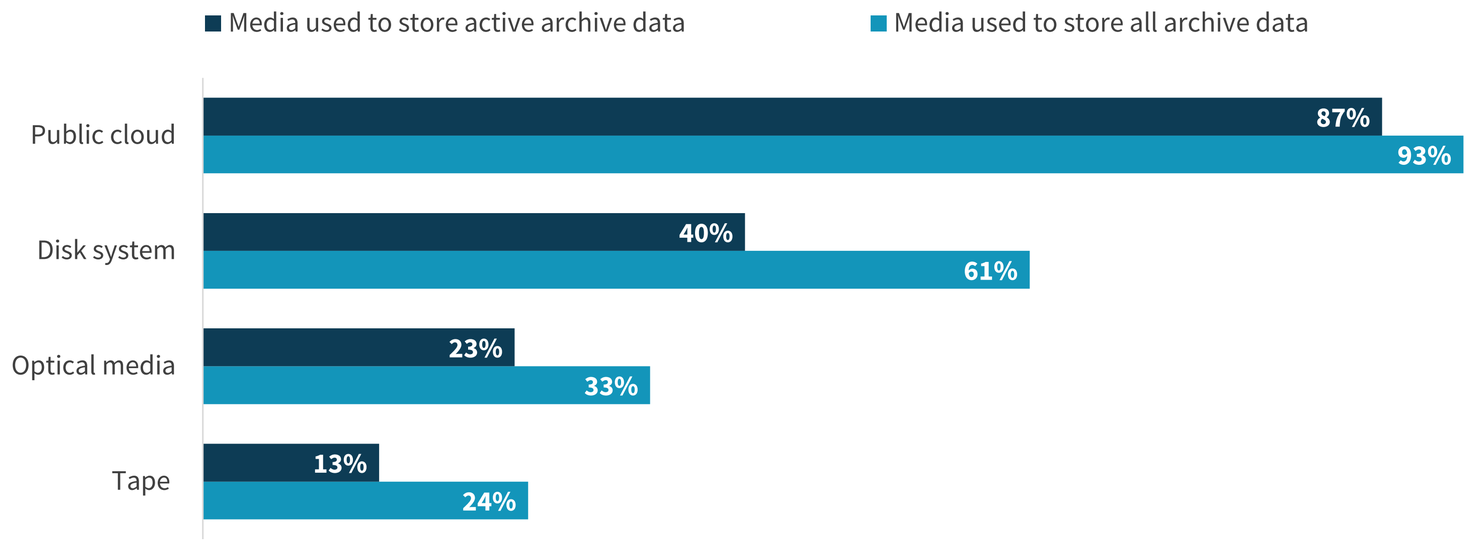
Cloud Archiving Is the New Norm

Public Cloud Is the Most Widely Used Storage Platform for Archive Data, but Most Consider Their Active Archive Topology to Be Hybrid

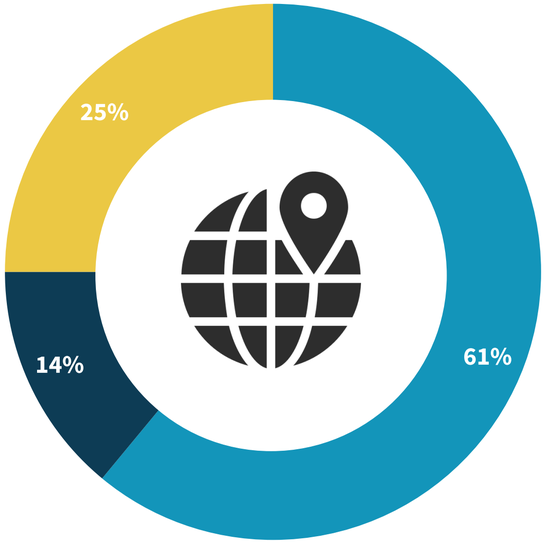
In terms of the media on which archive data is stored, an overwhelming number of organizations leverage public cloud services for both archive and active archive data. Indeed, 93% of organizations store their archive data in the cloud, while nearly nine in ten use these services for their active archive data.

From an infrastructure standpoint, active archive topologies are primarily hybrid, combining cloud and on-premises deployments. It should be noted that a quarter of respondents choose to deploy on primarily public cloud platforms, a vote of confidence for hyperscalers and service providers.

Media used to store archive and active archive data.



Active archive topology.



Back to Contents

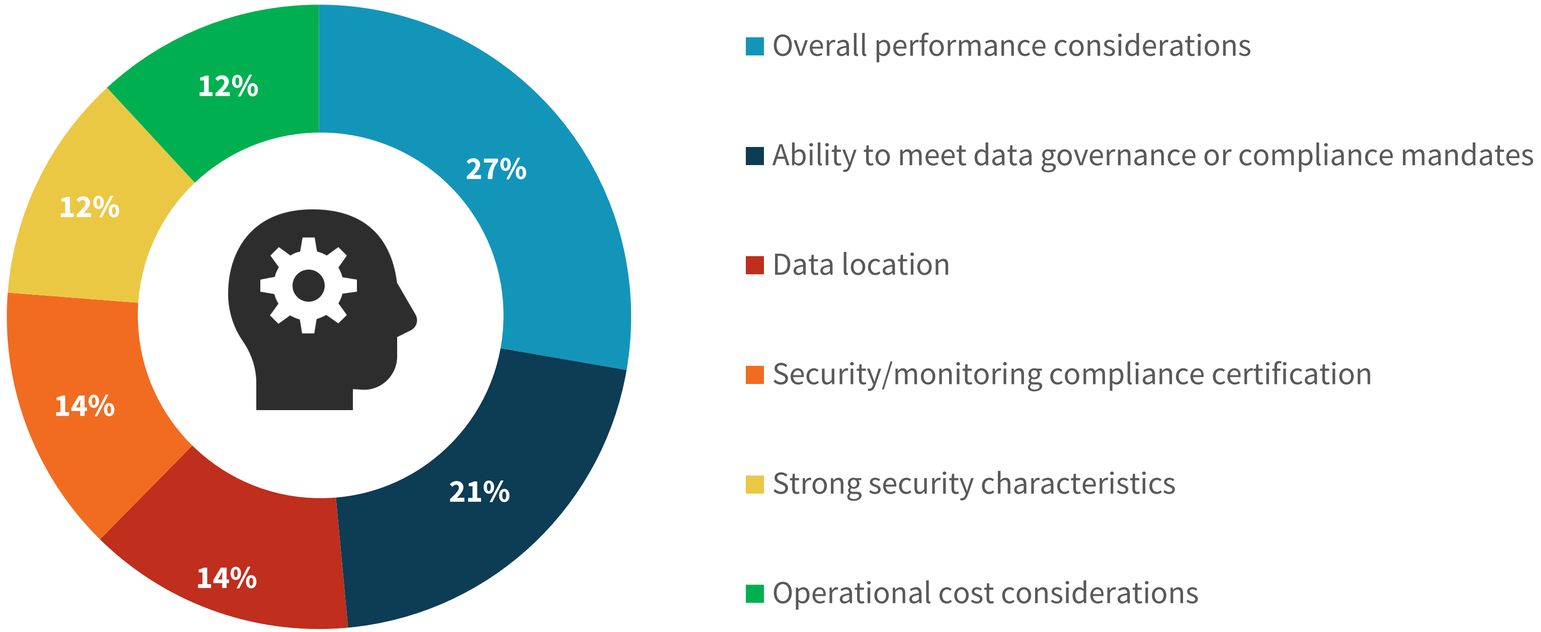
Performance Is Key to Topology Design, Followed by Compliance

Active archive topologies and solutions can vary based on a variety of factors: archive size, company size, primary use case, etc. While these are important in topology design, looking at decision considerations across the board and homing in on the technology characteristics, ESG identified that performance is at the top of the list, with more than a quarter of respondents citing it as the most significant consideration. Against a backdrop of digital transformation and massive data growth, it is not surprising to see performance (at scale) on top of the list. Afterall, it is a key factor in driving more efficiency and deriving accelerated returns on data asset investments.

Compliance and data governance mandates are also top of mind and drive decision-making processes for many organizations. The challenging and ever-changing regulatory compliance and privacy context is likely a root cause.

Surprisingly, cost is last on this list, with fewer than one in ten respondents identifying it as the most important factor. This is the sign of a more mature perception of the value of active archives: They are an investment for business optimization, not just a way to lower IT/data costs.

Most significant consideration in the selection of the topology used to support active archive.



Back to Contents

ABOUT SPECTRA LOGIC CORPORATION

Spectra Logic develops a full range of data storage and data management solutions that help customers store, manage, preserve and use vast amounts of data to advance their strategic missions and increase value for their organizations. Dedicated solely to data storage innovation for more than 40 years, Spectra Logic empowers organizations worldwide to harness the power of their data with scalable and modern solutions to accelerate breakthroughs and success in the market. To learn more, visit www.Spectralogic.com.

LEARN MORE

ABOUT ESG

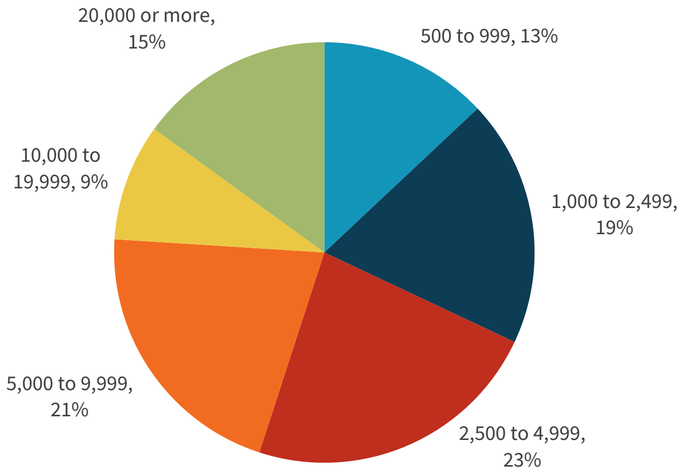
Enterprise Strategy Group is an integrated technology analysis, research, and strategy firm providing market intelligence, actionable insight, and go-to-market content services to the global technology community.

Research Methodology

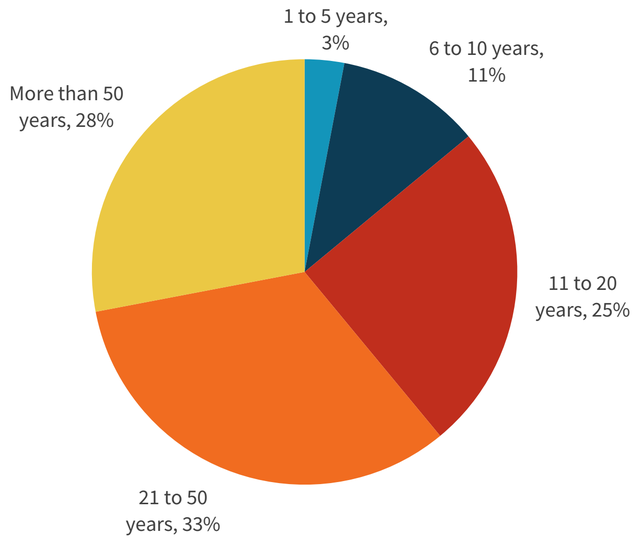
To gather data for this report, ESG conducted a comprehensive online survey of IT professionals from private- and public-sector organizations in North America (United States and Canada) between May 18, 2021 and May 25, 2021. To qualify for this survey, respondents were required to be IT professionals personally familiar with and/or responsible for data protection technology decisions, specifically data archiving and long-term retention strategies, for their organization. Respondents' organizations were required to be using an active archive technology solution. 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 150 IT professionals.

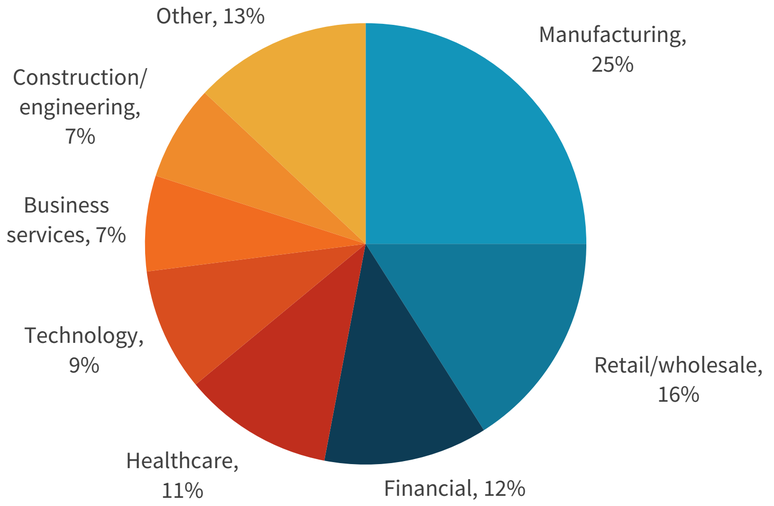
RESPONDENTS BY NUMBER OF EMPLOYEES



RESPONDENTS BY AGE OF COMPANY



RESPONDENTS BY INDUSTRY



Back to Contents

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.