October 2025 | White Paper Sponsored by Seagate Technology

Content Creation in the Age of Generative AI: Implications for Value and Scale

Adam Wright
Research Manager,
Global DataSphere and
Global StorageSphere, IDC

Greg Ireland
Senior Director,
Consumer Markets and GenAI for
Content Creators and Consumers, IDC

Edward Burns
Research Director,
Hard Disk Drive and Storage Technologies, IDC

THIS PDF USES HYPERLINKS

Table of Contents

CLICK ANY HEADING TO NAVIGATE DIRECTLY TO THAT PAGE.

In This White Paper

To better understand the impact of generative AI (GenAI) on content creation and data storage, Seagate Technology commissioned IDC to conduct a market research study. The study surveyed 1,000 business and IT decision-makers worldwide from large and midsize organizations across industries that use GenAI tools. IDC analysts also conducted qualitative research consisting of in-depth interviews with enterprise end users of generative AI tools and leveraged existing IDC syndicated research, such as IDC's Global DataSphere and StorageSphere, for perspectives on data generation, utilization, and storage.

It is important to note that IDC intentionally scoped this research to capture the enterprise demand-side perspective. The analysis does not encompass pure consumer markets or the enterprise supply-side of content creation technologies. As such, the findings should be considered a moderate view of generative AI's impact, and the true magnitude of content creation growth may ultimately prove even greater as these additional forces come into play.

Key Findings

By the time of this report's publication, GenAI has already begun to reshape the content creation landscape, enabling faster production, broader participation, and more sophisticated outputs across industries.

While the scale of impact varies, the trajectory points toward GenAI becoming a foundational element of modern content ecosystems. The study's findings also highlight an emerging challenge for enterprises: how to balance the expanding scale of content with the imperative to extract greater value from it.

Key findings include the following:

GenAI is increasing content volumes, richness, and complexity.

- 78.4% of respondents are now generating content that they didn't previously create, often because it was too costly or time-consuming to produce using traditional methods.
- Two-thirds of respondents (66.0%) indicate that — compared to before the adoption of generative AI tools — the total number of content files created (e.g., documents, images, videos, and presentations) has somewhat increased or significantly increased; more than a quarter (27.5%) say it has significantly increased.

- More than 75% of respondents indicate that video file content will increase by at least two times the current amount or more over the next five years due to the use of GenAI.
- More than 70% of respondents note that the average file size has already slightly or significantly increased as a result of using GenAI; 31% report that files are "significantly larger, richer, higher resolution, or more complex."
- 30.9% of respondents already see significantly larger file sizes due to GenAI, with another 41.8% seeing somewhat larger sizes.

GenAI is accelerating content replication.

- 56.9% of respondents report creating multiple variations of the same content more often since adopting GenAI.
- 45.8% say the ease of generation has led to their organization saving higher volumes of similar or excess files.
- 24.9% systematically store all variations, 36.1% selectively store the best versions, and 24.9% usually discard extras.

GenAI is sparking the demand for more and longer content retention.

- Only 34.2% of respondents say their storage infrastructure is fully optimized for GenAI, while 38.4% are only partially prepared.
- 66.1% of respondents expect moderate or significant storage growth over the next two years due to GenAI.
- 42% are planning for the need to retain data for longer periods by implementing data tiering and archiving strategies, including archiving older or less frequently accessed data to cost-effective storage tiers.

GenAI is democratizing content creation.

- 74.7% of respondents say GenAI enables employees outside of traditional creative roles to generate content.
- More than 42% report content creation is now significantly faster, with another 37% reporting it is somewhat faster.

Introduction

The world is generating and storing more data than ever before, and the pace is accelerating. IDC projects the creation of 213,557 exabytes (EB) of data globally in 2025, more than doubling to 527,469 exabytes by 2029.

To put that into perspective, a single exabyte equals one million terabytes, or one billion gigabytes. IDC expects the total installed base of storage capacity to grow from 11,243 exabytes in 2025 to 19,341 exabytes in 2029, and the total amount of data stored is projected to reach 13,986 exabytes by 2029.

Against this backdrop, GenAI has emerged as a key accelerator of digital content creation. GenAI is a type of AI that can create new content, such as text, images, music, video, or code, by learning from existing data. It uses models to generate outputs that are similar to the data it was trained on, enabling it to produce new content. In just two years, GenAI has moved from a nascent technology to a widely used tool. Enterprises are integrating it into workflows, creative platforms, and consumer applications at a rapid pace.

For consumers and prosumers, GenAI is changing the methods of creating and sharing text, images, and videos, thereby increasing the speed of creation and enabling entirely new forms of expression. Enterprises are building GenAI into marketing automation, customer support, code development, product design, and more, promising to deliver faster turnaround times, higher productivity, and measurable business results.

The results of the survey that informed this report and user interviews reflect this momentum: More than 50% of respondents have been using GenAI tools for over a year, and 45.5% of respondents report that daily use is common. By comparison, IDC's research shows approximately 69% of the world's population uses the internet, and about 58% use a smartphone. A further 40% of respondents have analyzed ROI and found a strong return, and nearly one-third (29.3%) say their organization now requires GenAI use as part of annual performance metrics.

This adoption wave is already changing how, and by how much, enterprises are creating content. While they don't store every AI-generated output, the overall effect on the digital ecosystem inevitably extends to data generation and storage demand.

GenAI holds the potential to have a significant impact on data generation and storage needs over the coming years in four primary ways:

Richer, more complex content:
GenAI leads to the production of larger files and more sophisticated media.

Greater replication and reuse:
GenAI has a multiplying effect on content generation, leading to more variations and derivative content.

Longer data retention:
GenAI is increasing the amount of data that enterprises are generating and saving, which may lead to longer periods of data storage.

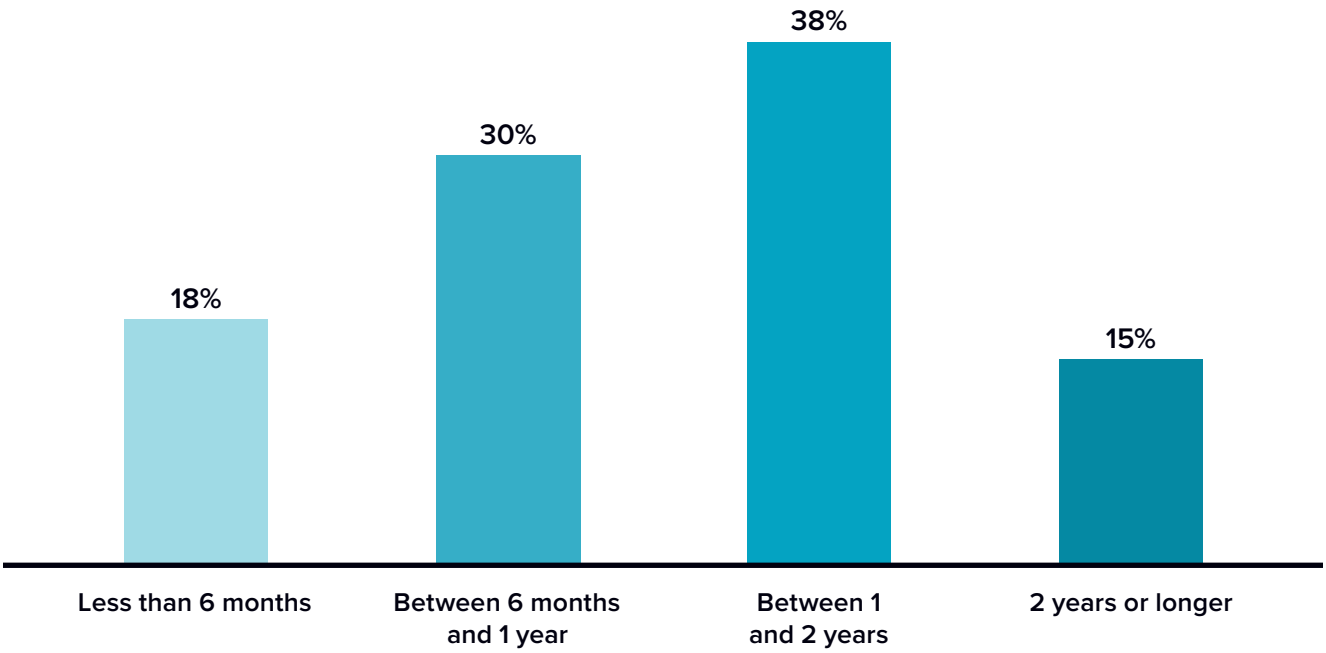
Democratization of content creation:
GenAI opens creation to more people in more roles, with fewer barriers.

While GenAI's long-term effect on stored data volumes is still emerging, its influence on the speed, scale, and accessibility of content creation is already clear. Over time, as richer formats proliferate and retention strategies evolve, these shifts are likely to ripple into higher storage demand, particularly in high-capacity tiers based on systems underpinned by hard disk drives (HDDs). Across AI training and AI inference workloads, high-capacity hard drive tiers can deliver the throughput and cost efficiency needed to support sustained model development and the persistent data sets that keep AI systems responsive at scale. There is also growing recognition that AI-generated content may hold increasing value over time, for immediate business outcomes and as potential fuel for future AI model training and analytics.

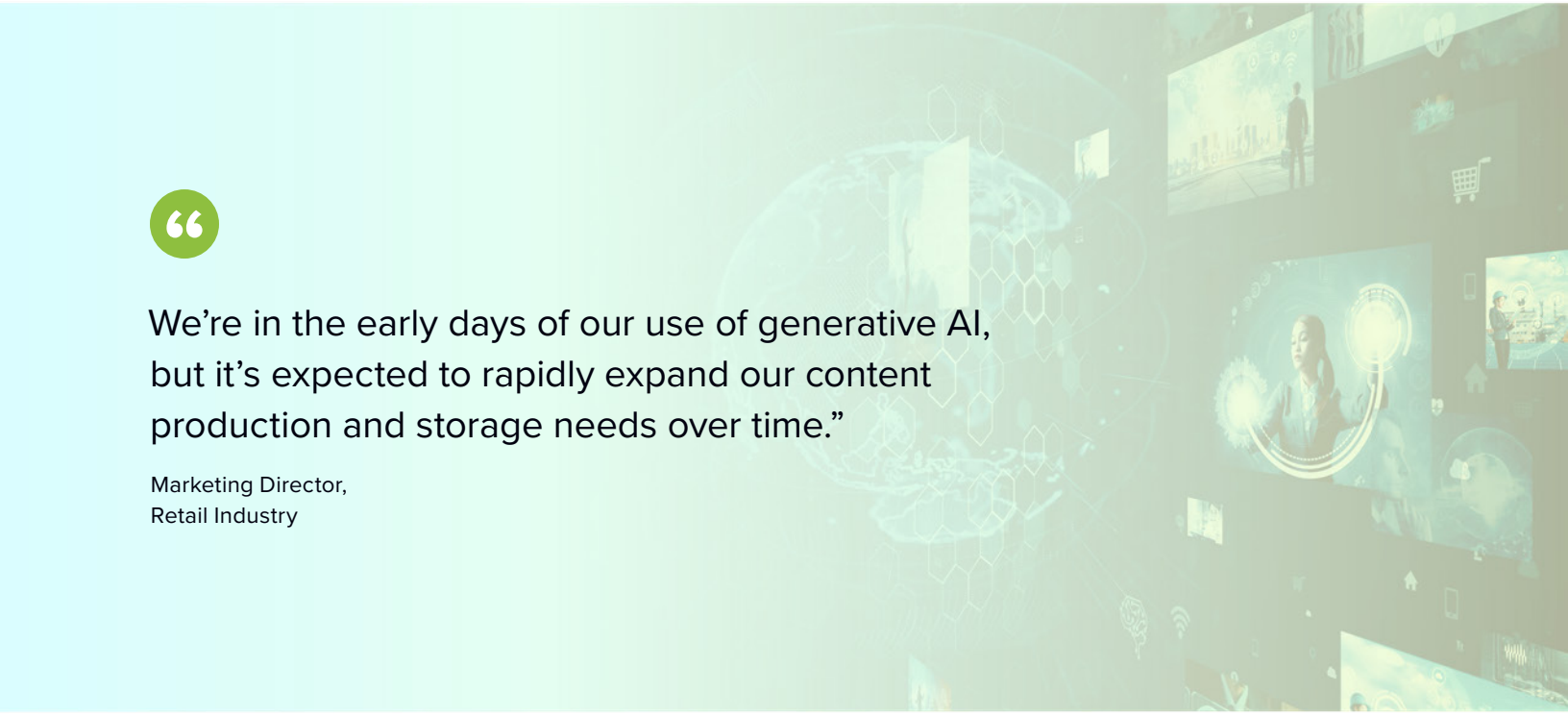
FIGURE 1

How long have you or your organization been using generative AI tools at work?

(Percent of respondents)



Note: The percentages add up to more than 100% due to rounding. n =1,034; Source: IDC's *Custom Survey 2025: Seagate GenAI Data Creation Survey*



"We're in the early days of our use of generative AI, but it's expected to rapidly expand our content production and storage needs over time."

Marketing Director,
Retail Industry

GenAI Use Cases

Generative AI adoption is expanding rapidly, with organizations applying the technology across a wide spectrum of business functions.

IDC's survey results show that usage is not only widespread but also frequent, with many organizations embedding GenAI into daily workflows. The breadth of use cases reflects the versatility of the technology, from creative media production to operational documentation, and highlights its growing role as a core enabler of modern business processes.

Marketing and sales content leads all reported use cases, with 71.6% of survey respondents indicating their organization uses GenAI to create ad copy, social media posts, pitch decks, and personalized content. In the retail industry, this can mean generating thousands of localized product descriptions, seasonal campaign visuals, and targeted email offers, each tailored to specific customer segments and often integrating AI-generated imagery or video. In finance, institutions are using GenAI to produce personalized investment recommendations, compliance-reviewed promotional brochures, and client outreach campaigns. These outputs often combine text, graphics, and multimedia, which can increase file sizes and their potential impact on storage, especially when multiple variations are created for different audiences.

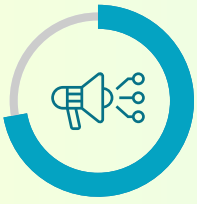
A majority of respondents (69.5%) cited business writing and documentation as the second most common use case. In healthcare, this includes drafting patient education materials, summarizing clinical trial data, and preparing insurance or

billing documentation, which is content that frequently requires secure, long-term retention for compliance. In manufacturing and industrial environments, GenAI produces maintenance manuals, safety procedure updates, and quality control reports. While text-based files are relatively small compared to rich media, their sheer frequency and the retention requirements tied to compliance, safety, or operational records can contribute significantly to long-term storage needs.

Creative media generation ranks third, with 58% of organizations using GenAI to produce images, videos, and presentation materials. In media and entertainment, this includes generating concept art, storyboards, B-roll footage, and visual effect elements for film, television, and streaming platforms, producing large files and multiple iterations that require storage for review, localization, or reuse. In retail, marketing teams use GenAI to create lifestyle photography, promotional videos, and ad campaign materials at a scale that would have been cost-prohibitive with traditional production methods. Each iteration or localized variant can add substantially to overall data volumes, particularly when kept for future repurposing.

Generative AI is also beginning to influence video surveillance by enabling advanced analytics and automated content generation from captured footage, ranging from security monitoring to tracking foot traffic in retail stores and patient flow in healthcare facilities. As these applications expand, the volume of high-resolution, retained video data is likely to grow, driving demand for scalable, cost-effective storage solutions capable of handling continuous, media-heavy workloads.

Reported use cases of GenAI in daily workflows:



72%
Marketing and sales content

70%
Business writing and documentation

58%
Creative media generation

GenAI Adoption Drivers

The drivers for adopting GenAI tools are as diverse as the use cases themselves:

Increasing efficiency and productivity (51.8%) is the top driver as organizations seek to reduce manual effort and speed up content creation cycles.

The second leading driver, **improving quality and consistency** (51.0%), reflects a desire to ensure that output meets brand guidelines, regulatory standards, or quality benchmarks across all content types.

The third key driver, **enhancing customer experience** (47.6%), often involves personalizing interactions at scale, whether that's a retailer providing customized product recommendations, a healthcare provider tailoring wellness advice, or a financial institution offering client-specific insights.

While human prompts trigger much of today's GenAI output, where a user specifies the request, machine-generated prompting is already being leveraged. IDC's survey finds that 61.2% of respondents report some level of machine-generated automation in the activation methods of GenAI tools. In healthcare, automated prompting can generate daily patient progress summaries or flag anomalies for clinician review based on connected device data. In media and entertainment, it can produce real-time social media updates, promotional materials, or highlight reels during live events. By enabling continuous or event-driven content generation without direct human input, these automated workflows have the potential to multiply output volumes rapidly and at scale.

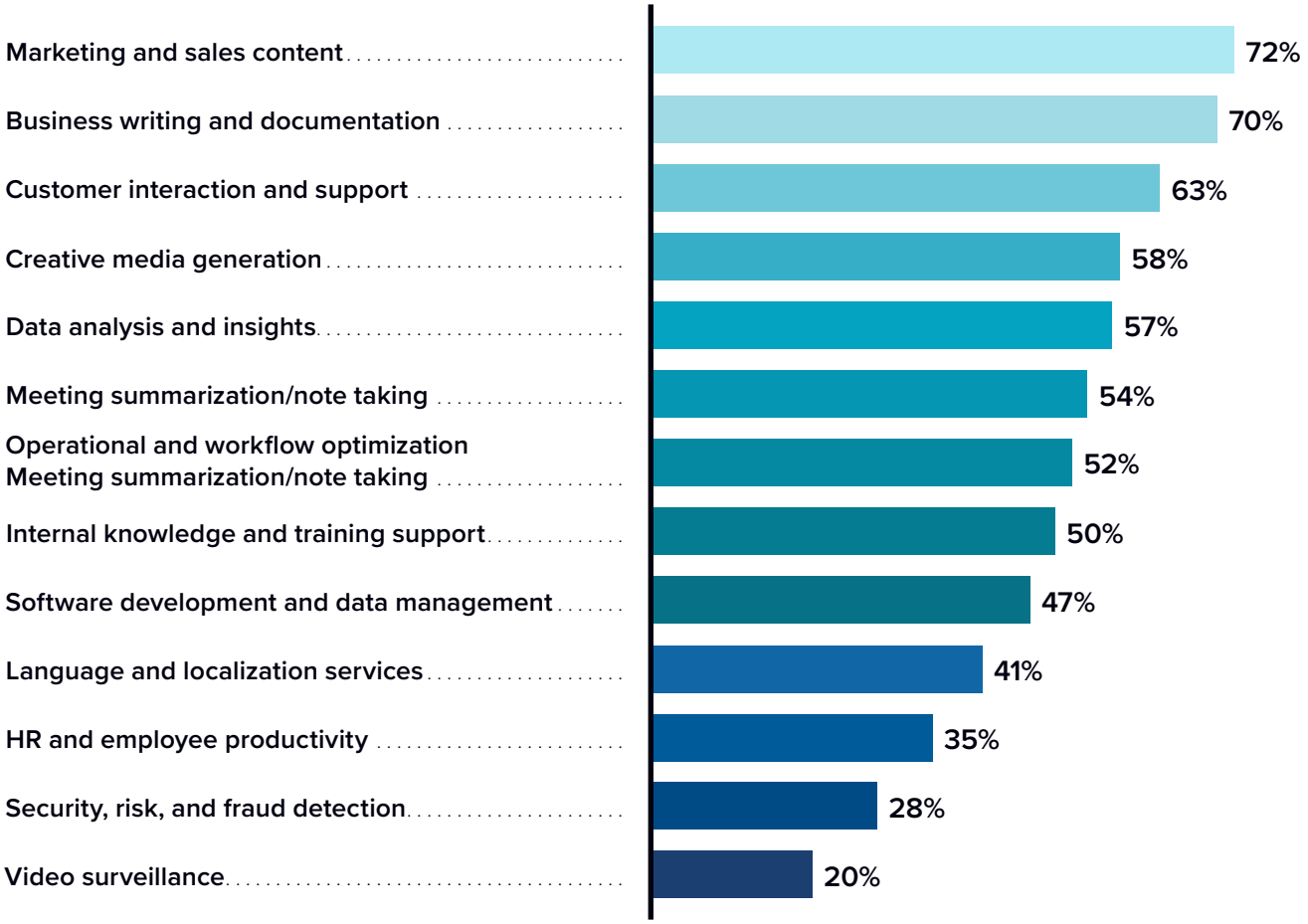
Consumer adoption is also emerging as a significant force, particularly in the creation of AI-generated images and videos for social media. Platforms such as TikTok, Instagram, and YouTube are hotspots for user-generated AI content, ranging from personalized avatars and stylized photos to short-form videos with AI-generated backdrops, effects, or scripts. While much of this content is transient from the end user's perspective, it often passes through the infrastructure of the hyperscale cloud

providers that host these platforms. Even ephemeral content can generate sizable short-term storage and processing loads, especially when users upload millions of AI-enhanced videos or images daily. In aggregate, this creates bursts of demand for hyperscaler storage, caching, and content delivery resources. Over time, as platforms increasingly retain AI-generated assets for analytics, personalization, or monetization purposes, consumer-driven GenAI activity could become a nontrivial contributor to long-term data storage growth in hyperscale environments.

FIGURE 2

For what purposes is your organization currently using generative AI?

(Percent of respondents)

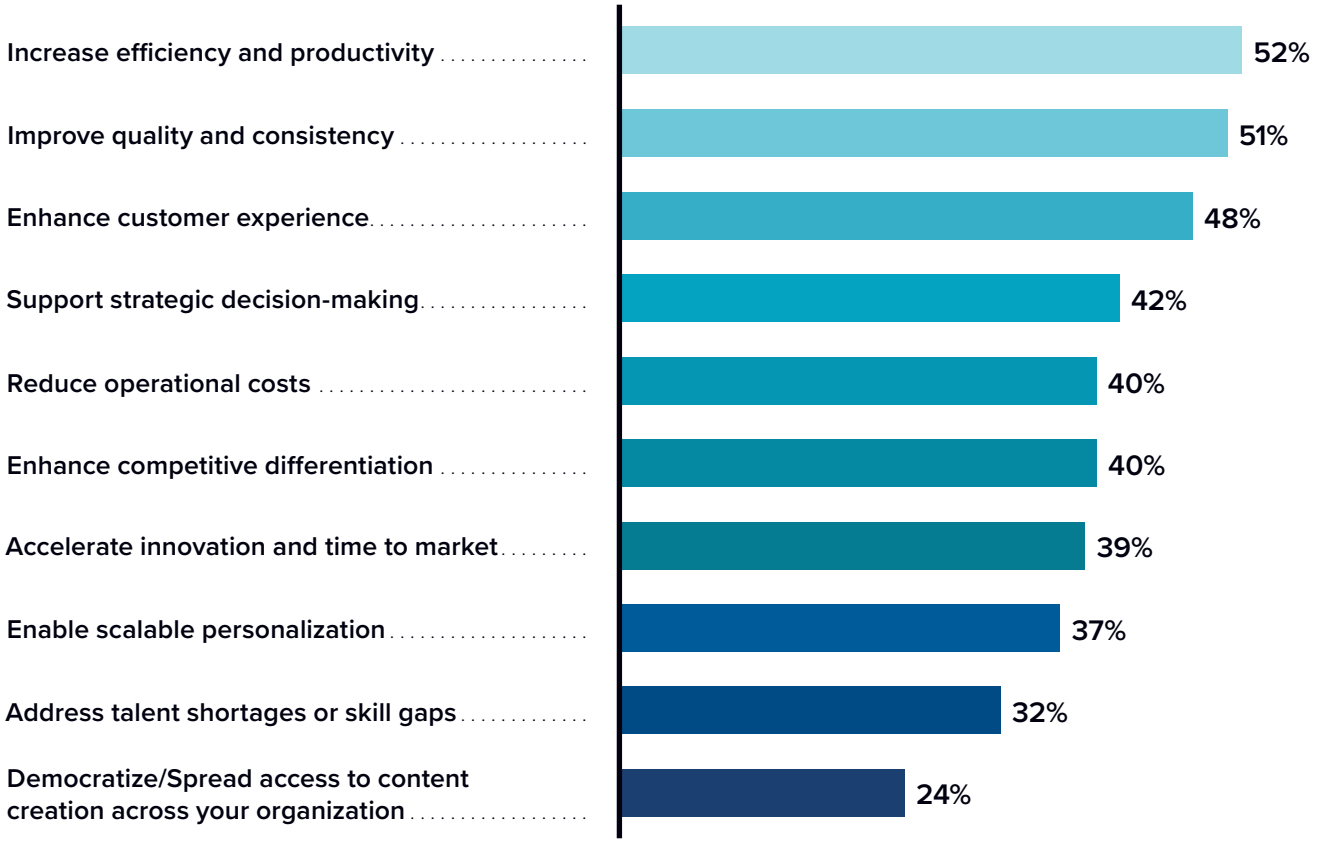


Note: Marketing and sales content includes: ad copy, social media posts, emails, landing pages, pitch decks, lead scoring, and personalization. Business writing and documentation include: reports, memos, emails, summaries, internal guides, and compliance updates. Customer interaction and support include: chatbots, virtual agents, and customer service automation. Creative media generation includes: images, videos, presentations, and scripts. Data analysis and insights include: customer feedback analysis, dashboards, trends, and predictive modeling. Meeting summarization/note taking includes: automatically generating summaries or action items from recorded meetings or live transcripts. Operational and workflow optimization includes: task tracking, meeting summaries, and process alerts. Internal knowledge and training support includes: onboarding materials, policy documentation, and internal communications. Software development and data management include: code assistance, debugging, data cleansing, and synthetic data creation. Language and localization services include: translation and multilingual content. HR and employee productivity includes: hiring support, performance reviews, and calendar/task tools. Security, risk, and fraud detection includes: cyberthreat summaries, compliance monitoring, and surveillance analytics. Video surveillance includes: analyzing footage for unusual behavior or specific trends, or identifying people/objects. n = 1,034; Source: IDC's *Custom Survey 2025: Seagate GenAI Data Creation Survey*

FIGURE 3

Which of the following best describes the primary drivers behind your organization's use of generative AI tools?

(Percent of respondents)

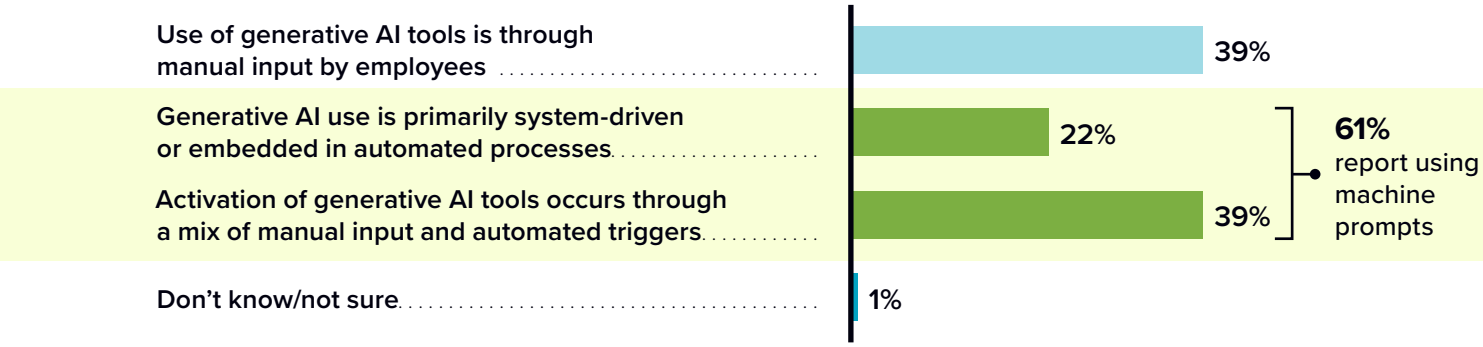


n = 1,034; Source: IDC's *Custom Survey 2025: Seagate GenAI Data Creation Survey*

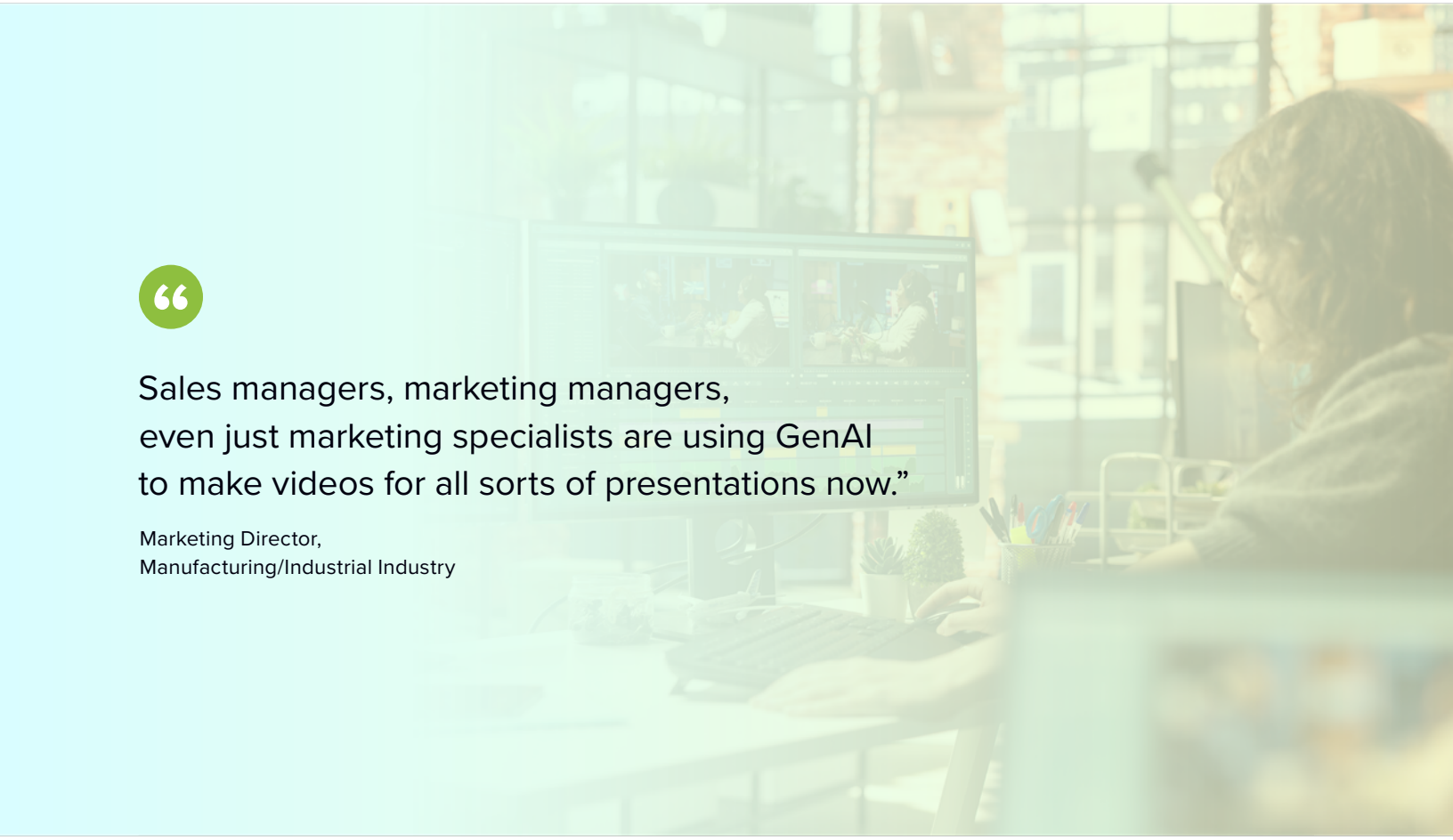
FIGURE 4

Which of the following best describes how generative AI tools are triggered or activated within your organization?

(Percent of respondents)



Notes: The percentages add up to more than 100% due to rounding. Use of generative AI tools is through manual input by employees, for example, only human prompts. Generative AI use is primarily system-driven or embedded in automated processes, for example, mostly through automated or machine-generated prompts. Activation of generative AI tools occurs through a mix of manual input and automated triggers, for example, both human- and system-generated prompts. n = 1,034; Source: IDC's *Custom Survey 2025: Seagate GenAI Data Creation Survey*



> Sales managers, marketing managers, even just marketing specialists are using GenAI to make videos for all sorts of presentations now."
>
> Marketing Director,
> Manufacturing/Industrial Industry

GenAI and Content Volumes, Richness, and Complexity

One of the most important dynamics at play is the creation of content that would not have existed without GenAI.

Survey results show that 78.4% of respondents are now generating content that they didn't previously create, often because it was too costly or time-consuming to produce using traditional methods. This net-new content can range from tailored product imagery for ecommerce listings to explainer videos for niche customer segments to simulation outputs in industrial training programs. Each represents an expansion of the organization's content footprint rather than a one-for-one replacement of manually created materials.

GenAI is also leaving a perceived mark on the quantity and quality of digital content that enterprises are producing. Survey results show that, even in these early days, 27.5% of respondents have seen a significant increase in content volume due to GenAI adoption, with another 38.5% reporting a moderate increase (for a combined 66% of respondents seeing an increase in data volumes already despite nascent adoption). While text remains the leading type of generated content, images and videos feature prominently, both of which carry a disproportionately large impact on data volumes due to their bigger file sizes.

In relation to video use cases, such as video surveillance and computer vision, GenAI is enhancing applications well beyond traditional security. In retail, it can analyze video to map foot traffic, track shopper behavior, and produce annotated clips, heat maps, and summaries for optimizing store layouts. In healthcare, it supports patient monitoring by detecting falls or unusual activity, generating event-triggered alerts, structured logs, and illustrative clips. In manufacturing, it interprets and augments assembly-line footage with overlays, visual markers, and AI-generated inspection reports that blend images, diagrams, and text. By automatically detecting, classifying, and contextualizing events in video streams, GenAI helps produce derivative content, such as tagged highlight reels, synthetic simulation footage, and multimodal records. This is all material that can

substantially increase data volumes due to the retention of both high-resolution source video and derivatives. However, GenAI can also transform continuous raw footage into compact event-based segments, metadata summaries, or vector embeddings, preserving analytical value while lowering file sizes.

GenAI is not only increasing the number of files produced but is also broadening the mix of formats. In many organizations, the most noticeable shift is in the creation of media-rich assets, such as high-resolution images, complex infographics, and enhanced video content. These formats tend to be significantly larger in size than plain text documents, meaning even small increases in output can become "needle movers" for data growth.

Survey results underscore this trend:

A majority of respondents (66.0%) expect the total number of files their organization creates to increase over the next five years as a direct result of GenAI.

More than 75% of respondents indicate that video file content will increase by at least two times the current amount or more over the next five years due to the use of GenAI.

Beyond sheer quantity, GenAI is contributing to larger, richer, and more complex files.

More than 70% of respondents note that the average file size has already slightly or significantly increased as a result of using GenAI; 31% report that files are "significantly larger, richer, higher resolution, or more complex."

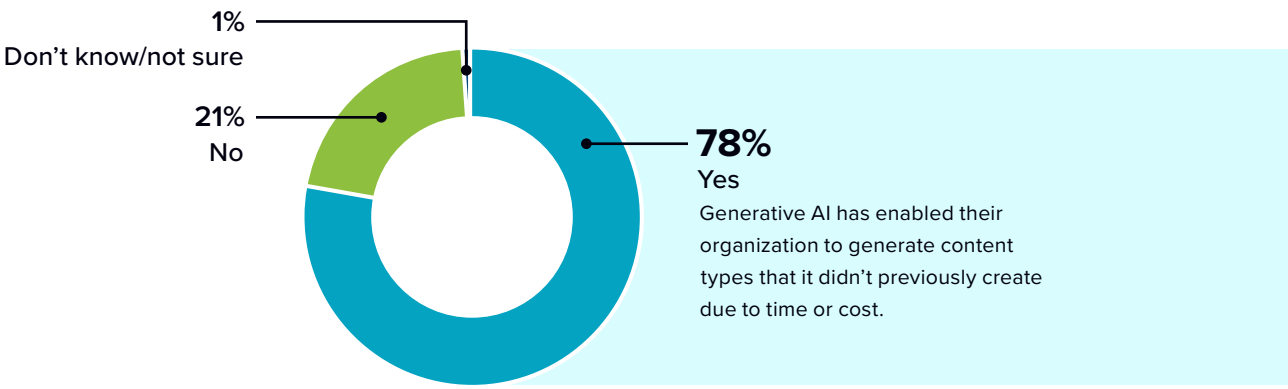
This is often the result of higher resolutions, more detailed graphics, layered image compositions, or video with enhanced frame rates and effects. Even for text-based content, richer formatting, embedded images, and multimedia integration can add considerably to storage requirements. From a storage perspective, this shift matters because larger average file sizes amplify the impact of volume growth. A modest percentage increase in file size, when multiplied across millions of new files, translates into material increases in data storage needs.

When viewed together, trends such as rising content volumes, richer formats, larger file sizes, and more net-new material paint a clear picture of GenAI as a driver of incremental data growth. While enterprises will not store all generated content long term, the increases in media-rich output and average file sizes make the storage implications more pronounced.

FIGURE 5

Has generative AI enabled your organization to generate content types that it didn't previously create due to time or cost?

(Percent of respondents)

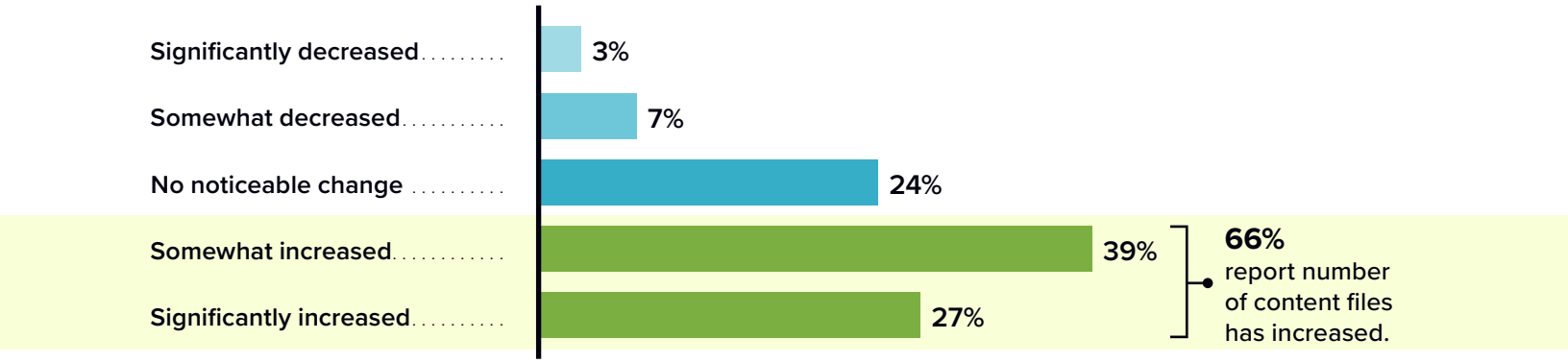


n = 1,034; Source: IDC's *Custom Survey 2025: Seagate GenAI Data Creation Survey*

FIGURE 6

Compared to before your organization adopted generative AI tools, how has the total number of content files created (e.g., documents, images, videos, and presentations) changed due to the use of generative AI?

(Percent of respondents)

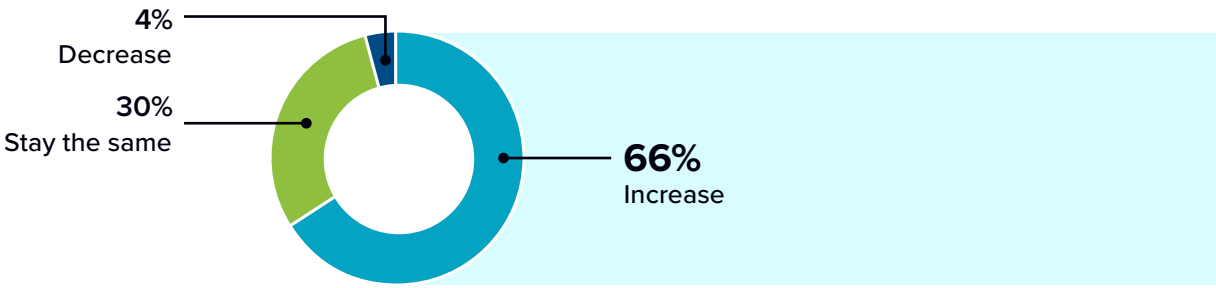


n = 1,034; Source: IDC's *Custom Survey 2025: Seagate GenAI Data Creation Survey*

FIGURE 7

Do you expect the number of files or content created in your organization to increase, decrease, or stay the same over the next five years due specifically to generative AI?

(Percent of respondents)

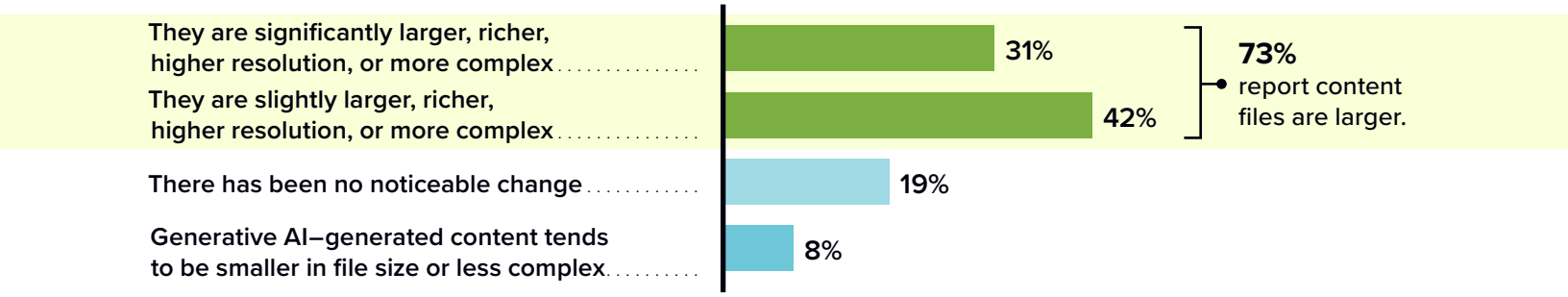


n = 1,034; Source: IDC's *Custom Survey 2025: Seagate GenAI Data Creation Survey*

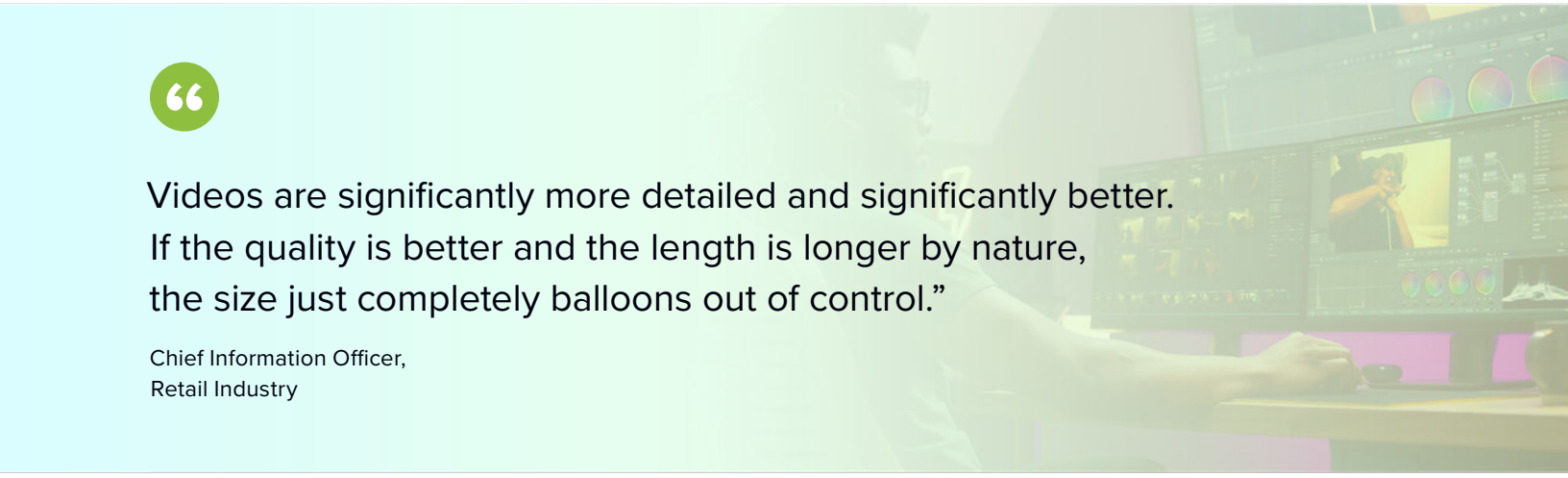
FIGURE 8

How has the use of generative AI impacted the average file size of content?

(Percent of respondents)



n = 1,034; Source: IDC's *Custom Survey 2025: Seagate GenAI Data Creation Survey*



"Videos are significantly more detailed and significantly better. If the quality is better and the length is longer by nature, the size just completely balloons out of control."

Chief Information Officer,
Retail Industry

GenAI and Content Replication

GenAI influences not only content creation but also the methods of content replication, storage, and reuse.

A defining characteristic of AI-assisted creation is the ease with which multiple variations of a single asset can be produced, whether to test different messaging, explore alternative designs, or refine creative direction.

IDC's research finds that:

56.9% of respondents are creating multiple variations or versions of the same content more often since adopting GenAI tools.

45.8% say the ease of content generation has led to them saving higher volumes of similar or excess files.

Not all variations that GenAI generates are intended to persist, however. Some are purely ephemeral, used for quick experimentation, internal review, or iterative prompt refinement. Others become part of a growing library of saved content, either for immediate use or for potential future application. This dual nature of AI-generated variations introduces new questions around storage, management, and retrieval.

The survey results reveal diverse organizational approaches to managing these variations:

24.9% systematically store and manage all variations.

36.1% selectively store only the best versions.

24.9% usually discard extra variations.

Version control and deduplication strategies are also common. Nearly half of respondents (45.0%) report that their organization actively uses these practices to manage GenAI outputs, and another quarter (24.8%) plan to implement them in the future.

For some organizations, saving additional variations can be a strategic choice and have latent value. They can reuse these assets or adapt them for future campaigns, serving as input for model retraining and fine-tuning, or providing inspiration and building blocks for new projects. In AI/ML contexts, a large, diverse set of variations can enhance inferencing, speed up the creation of related content, and even support analytics-driven decision-making. In industries such as media and entertainment, preserved variations can help rapidly assemble alternative edits or region-specific versions of content; in retail, they can provide a deep bench of product imagery for personalization and seasonal promotions. These examples highlight the growing value of data as organizations increasingly recognize its potential to drive innovation, improve efficiency, and unlock new opportunities across various industries and applications.

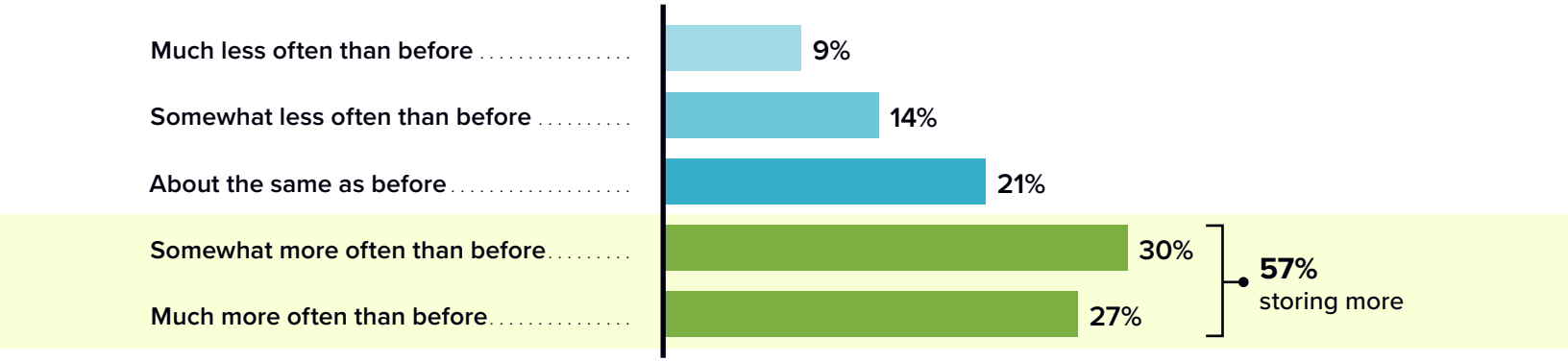
While these benefits are attractive, it is important to note that not every organization is ready to embrace the broad-scale saving of replicated content. In some cases, the ROI remains unclear, particularly if they only use variations once. Storing every output also has downstream implications for storage capacity, search and retrieval systems, and data governance. Budget constraints and competing IT priorities can further limit investment in the infrastructure needed to manage large volumes of AI-generated content. For these reasons, some business cases may not be clear, with decision-makers preferring to store only what is demonstrably valuable rather than what might have potential future use.

Overall, GenAI's ability to rapidly produce variations is changing how organizations think about content creation, experimentation, and reuse. The balance between ephemeral iteration and strategic retention will play an increasingly important role in shaping how these tools integrate into creative and operational workflows.

FIGURE 9

Since adopting generative AI tools, how often are you creating multiple variations or versions of the same content (e.g., different text drafts, image variations, and video edits)?

(Percent of respondents)

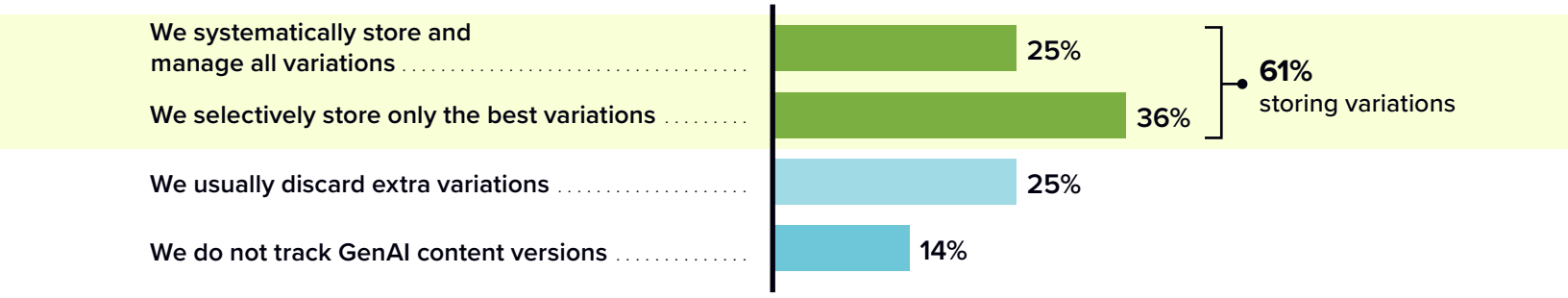


Note: These percentages add up to more than 100% due to rounding. n = 1,034; Source: IDC's *Custom Survey 2025: Seagate GenAI Data Creation Survey*

FIGURE 10

Which best describes your organization's approach to managing content variations that generative AI produces?

(Percent of respondents)

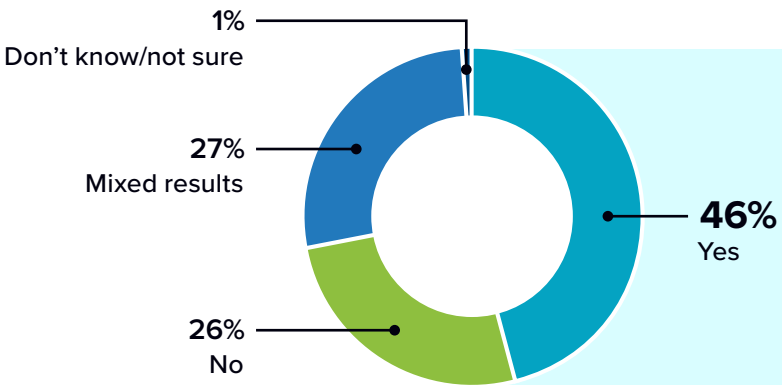


n = 1,034; Source: IDC's *Custom Survey 2025: Seagate GenAI Data Creation Survey*

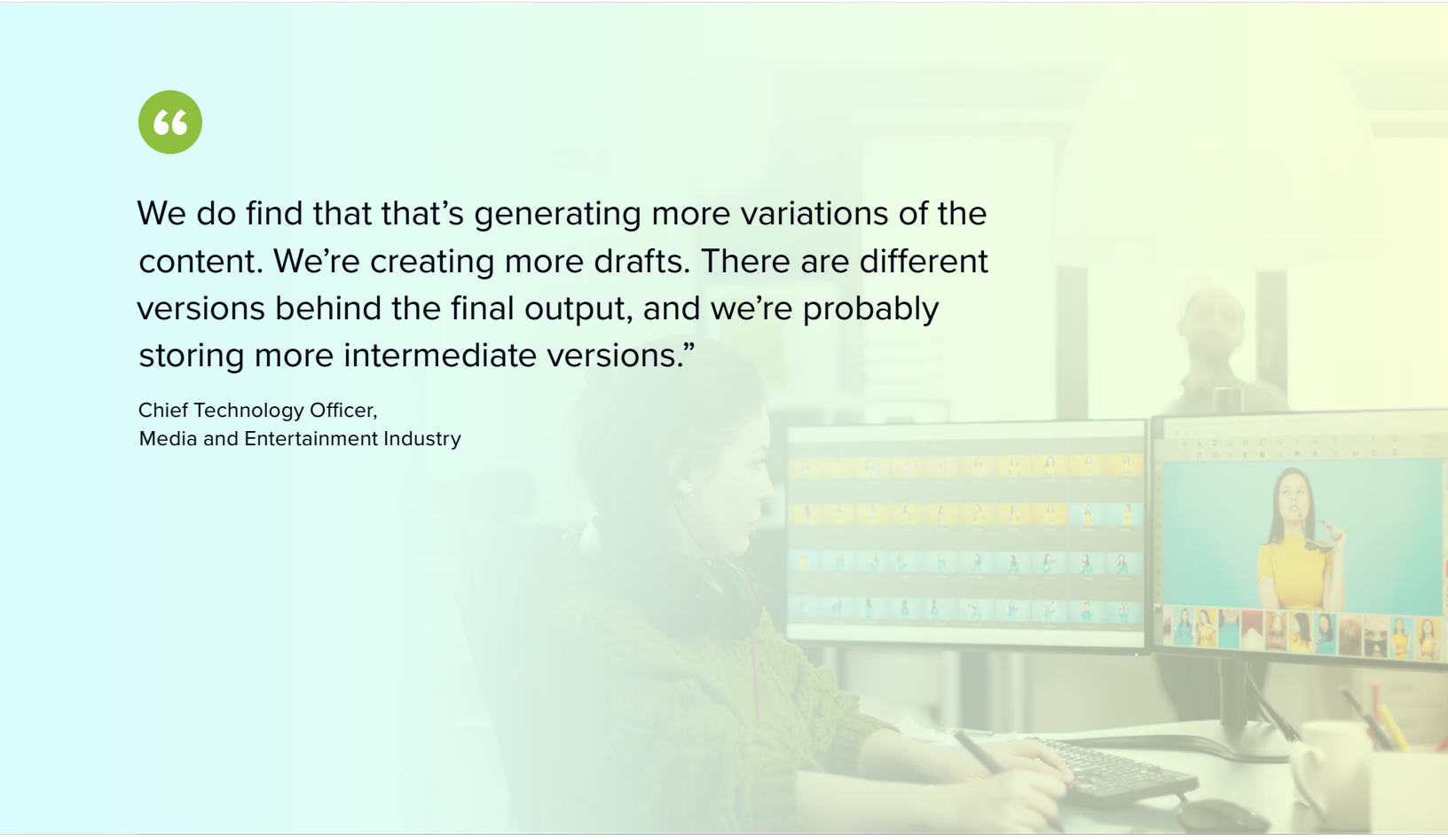
FIGURE 11

Has the ease of generating content via generative AI led to your organization saving higher volumes of similar or excess files?

(Percent of respondents)



n = 1,034; Source: IDC's *Custom Survey 2025: Seagate GenAI Data Creation Survey*



"We do find that that's generating more variations of the content. We're creating more drafts. There are different versions behind the final output, and we're probably storing more intermediate versions."

Chief Technology Officer,
Media and Entertainment Industry

GenAI and Content Retention

GenAI is not only transforming the methods of content production but is also beginning to influence the length and quantity of content retention.

IDC's research finds early signs that storage requirements are beginning to shift. The ease of generating new outputs, coupled with the potential latent value of AI-created content, is prompting many organizations to reconsider their storage strategies.

Enterprises are already feeling the early impacts on storage infrastructure. To handle the expected surge in generated data due to GenAI in the coming years, organizations are planning to implement an eclectic range of storage strategies.

Survey results show that:

The introduction of GenAI tools has led 65.1% of respondents to upgrade or expand their storage capacity, with an additional 24.9% expecting to do so in the near future.

42.5% of respondent organizations are planning for the need to retain data for longer periods by implementing data tiering and archiving strategies, including archiving older or less frequently accessed data to cost-effective storage tiers.

Only 34.2% of respondents say they have fully optimized their current storage infrastructure to handle GenAI-generated content, while another 38.4% are only partially prepared.

These adjustments reflect the tangible, short-term storage impacts as content volumes grow, not just from net-new content but also from the accumulation of variations, higher-resolution media, and richer file formats.

GenAI's influence is also visible in governance and life-cycle policies. According to the survey:

59.1% are still in the process of updating or are yet to determine if they need to make changes, which suggests that many are being caught off guard, reacting to storage and governance challenges rather than proactively planning for them.

39.0% of organizations have implemented new data governance or content life-cycle policies specifically in response to GenAI adoption.

Anticipated increases in retention and storage demand are also shaping strategic planning:

66.1% of respondents expect a moderate or significant increase in storage needs over the next two years due to GenAI.

Strategies to accommodate this growth vary: 50.7% of respondents intend to expand on-premises capacity, 48.7% plan to grow cloud storage footprints, and 46.3% expect to implement data compression and deduplication policies.

Hyperscalers, who host and manage massive amounts of consumer- and enterprise-generated content, are also a key part of this retention story. From social media images and short-form videos to user-generated AI-enhanced media, a large portion of this content flows through hyperscale infrastructure. GenAI tools embedded in consumer platforms make it easier to create and post richer media, which can lead to an increase in the volume of assets that hyperscalers store and, indirectly, drive growth in the HDD-intensive storage that underpins these environments. The large number of respondents indicating that they plan to implement storage tiering and archiving is a positive indicator for HDD demand because it not only remains the most economical storage technology on a cost-per-terabyte basis but is also the most durable solution for long-term storage, supporting the majority of enterprise workloads with optimal capacity and optimal performance.

Given the potential future value of AI-generated content, organizations might choose to retain GenAI-generated content longer than traditionally produced material. They can retain outputs for reuse, repurpose, or adapt them for future campaigns or projects. These outputs can also serve as data sets for improving AI models through training and fine-tuning. Having a library of past AI-generated

content can reduce turnaround time for similar future requests, while retained content can also feed into other AI/ML workflows for analysis, personalization, or automation. Each of these could become compelling storage accelerators to watch in the coming years.

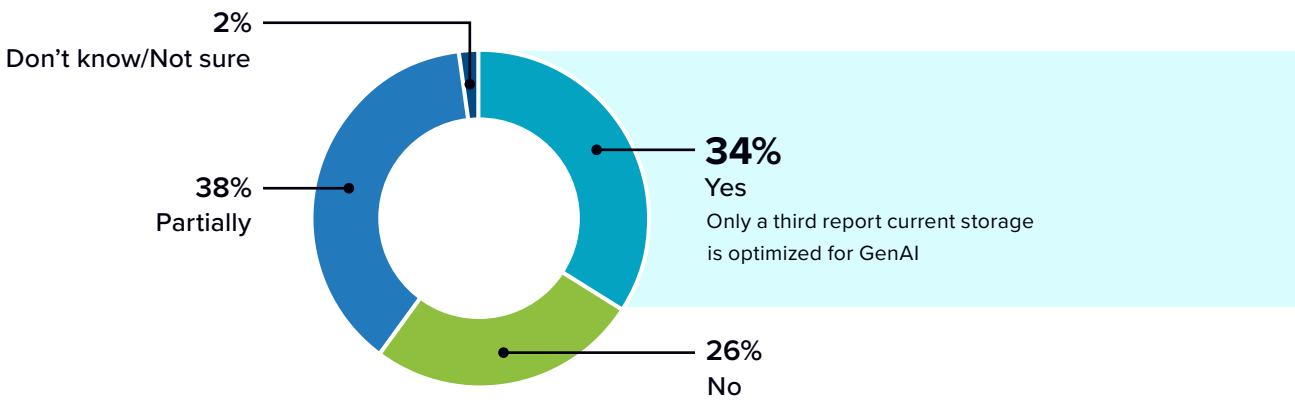
However, there are also countervailing pressures that could limit how much additional content organizations retain and for how long. Many AI-generated assets are ephemeral, serving short-lived purposes such as rapid prototyping or social media trends with a limited shelf life. Budget constraints, storage optimization goals, and unclear ROI can affect decision-making. Organization leaders need to establish a clear business case to invest in the storage and compute infrastructure required to house more content with speculative future value.

Ultimately, the balance between these drivers and constraints will shape how significantly GenAI impacts content retention in the coming years. While early indications point to moderate upward pressure on storage demand, especially in hyperscale and content-heavy sectors, much will depend on evolving governance strategies, economic conditions, and the proven ability to derive long-term value from retained AI-generated content.

FIGURE 12

Is your current storage infrastructure optimized to handle generative AI-generated content?

(Percent of respondents)

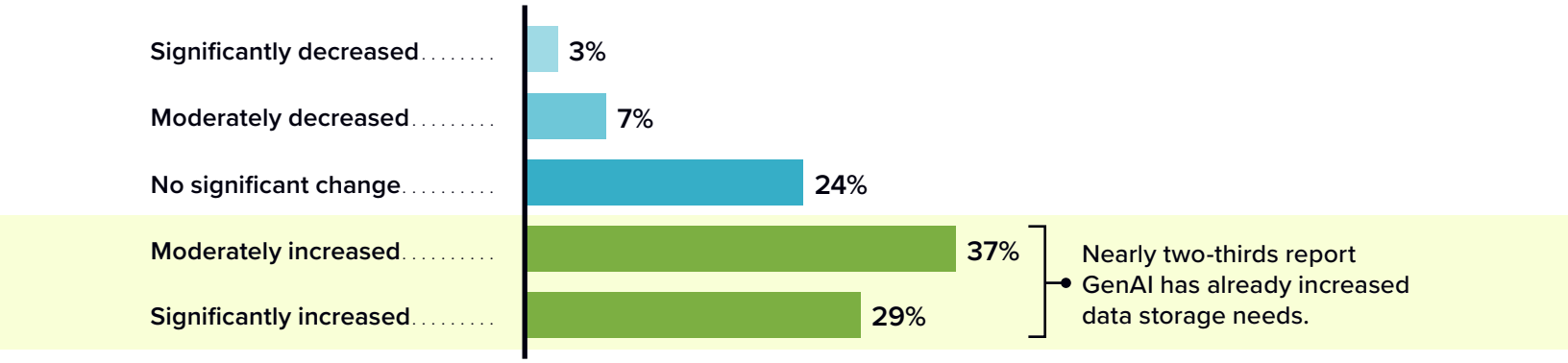


n =1,034; Source: IDC's *Custom Survey 2025: Seagate GenAI Data Creation Survey*

FIGURE 13

In the past 12 months, how has the introduction of generative AI tools affected your organization's data storage needs?

(Percent of respondents)

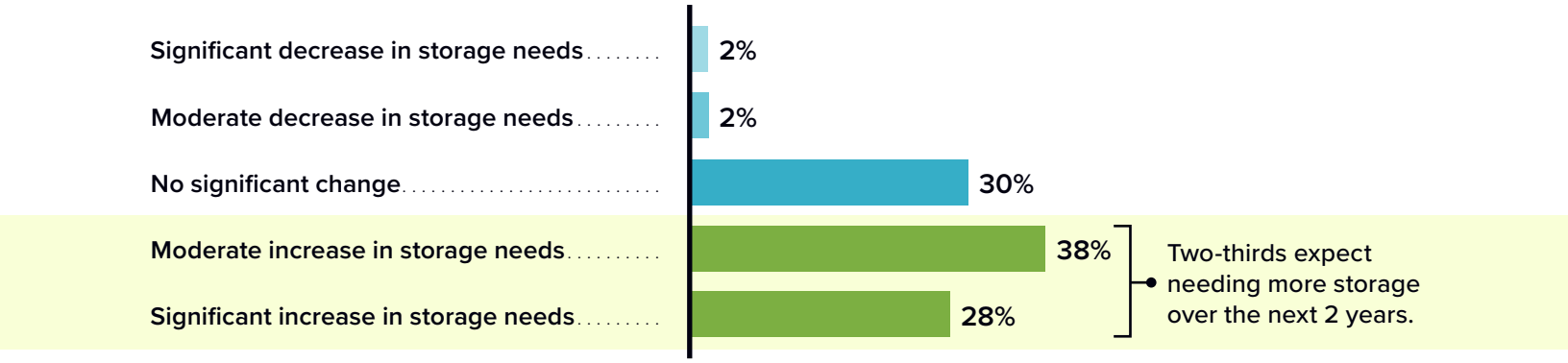


n = 1,034; Source: IDC's *Custom Survey 2025: Seagate GenAI Data Creation Survey*

FIGURE 14

How do you anticipate your organization's data storage needs changing over the next two years due to generative AI?

(Percent of respondents)

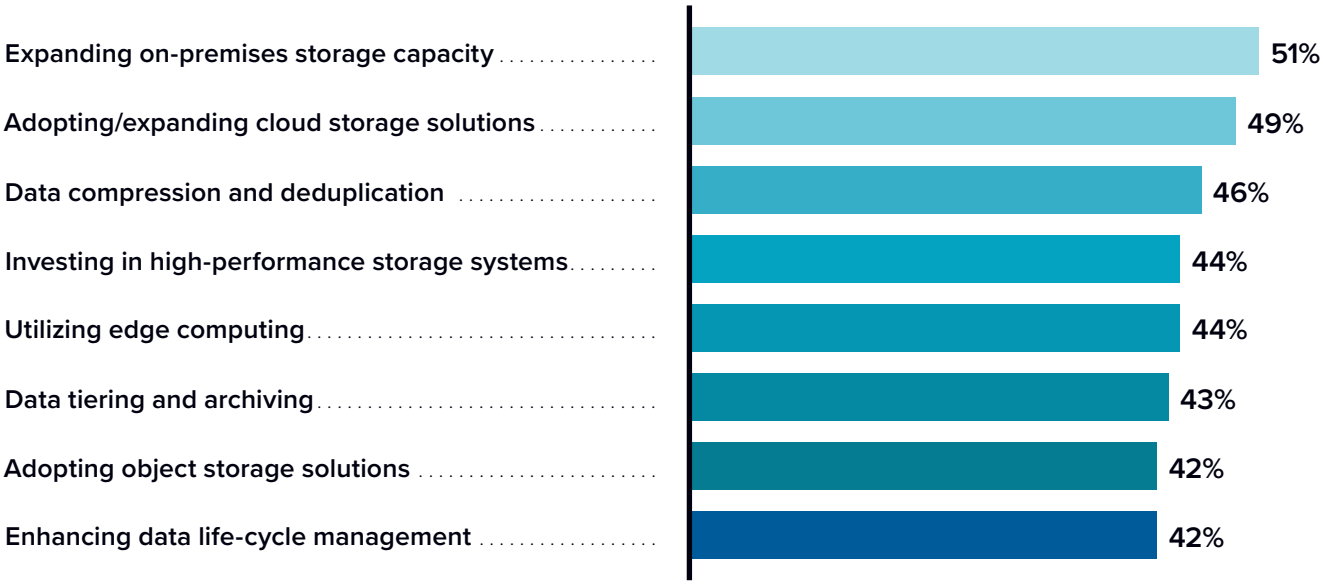


n = 1,034; Source: IDC's *Custom Survey 2025: Seagate GenAI Data Creation Survey*

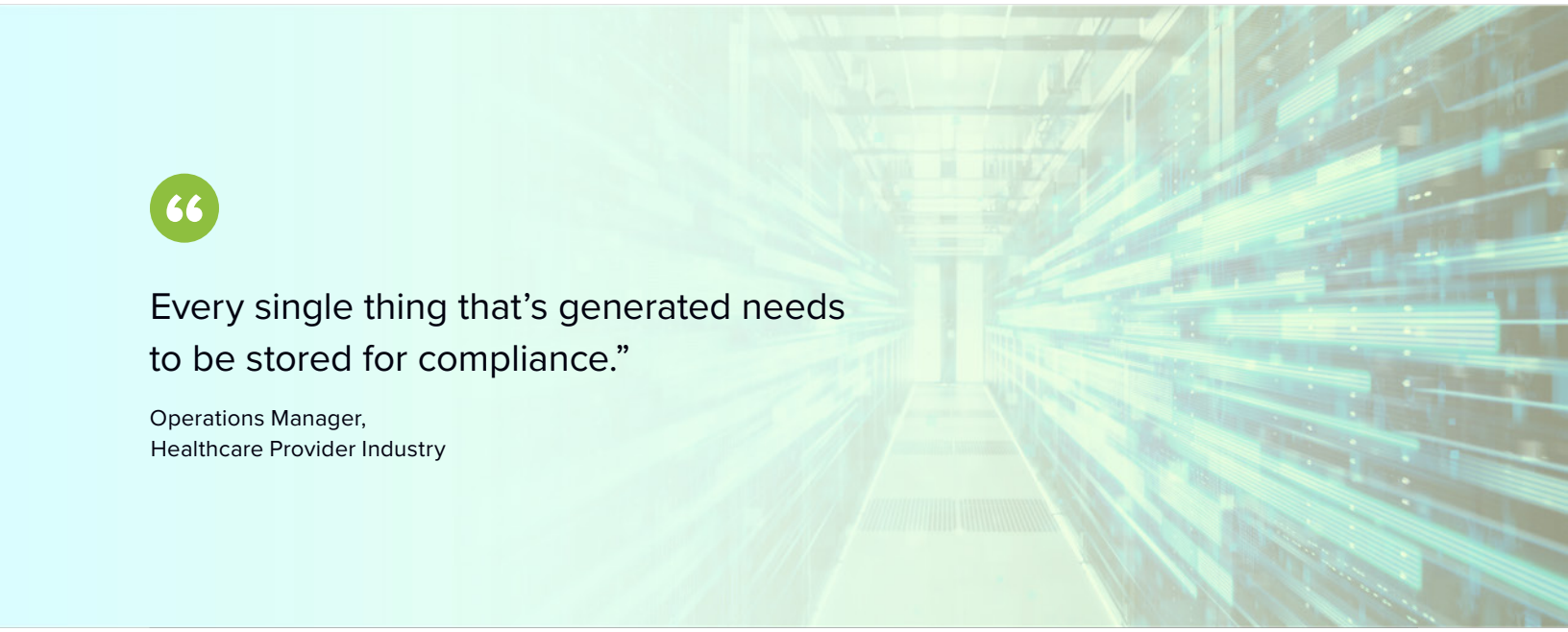
FIGURE 15

What storage strategies is your organization planning to implement to handle the increase in data that generative AI tools create?

(Percent of respondents)



Notes: Enhancing data life-cycle management: using automation tools to manage data movement and retention policies. Adopting/expanding cloud storage solutions: using public cloud storage services, hybrid cloud solutions. Data compression and deduplication: using data compression techniques to reduce storage footprint, implementing deduplication to eliminate redundant data. Investing in high-performance storage systems: using high-speed SSDs or NVMe storage for faster data access and processing. Utilizing edge computing: to process and store data closer to the source, reducing the need to transfer large volumes of data to centralized storage. Data tiering and archiving: implementing tiered storage strategies to optimize cost and performance, archiving older or less frequently accessed data to cost-effective storage tiers. Adopting object storage solutions: implementing object storage systems for handling large, unstructured data sets, using systems such as Amazon S3, Google Cloud Storage. Enhancing data life-cycle management: using automation tools to manage data movement and retention policies. n = 1,034; Source: IDC's *Custom Survey 2025: Seagate GenAI Data Creation Survey*



> Every single thing that's generated needs to be stored for compliance."

Operations Manager,
Healthcare Provider Industry

GenAI and the Democratization of Content Creation

One of the most significant ways GenAI is transforming the content landscape is by democratizing content creation: putting the ability to produce professional-quality text, images, videos, and other media into the hands of employees well beyond traditional creative or marketing teams.

IDC's research strongly supports this hypothesis, revealing that GenAI is lowering barriers to entry and enabling a wider range of people and roles to contribute to the creation process.

Survey results show that 74.7% of respondents indicate that GenAI allows employees outside of traditional creative roles to generate content. This shift expands the pool of content contributors, enabling organizations to tailor materials to more audiences and purposes. Reduced reliance on specialized content teams reinforces this trend, with 34.7% of respondents reporting that GenAI has lessened the need for dedicated creative resources for some content tasks. Creative professionals are not disappearing but evolving, focusing more on oversight, quality assurance, and strategic direction, while a broader group of employees is increasingly producing first drafts and campaign-ready content.

The impact on speed is also clear. More than 42% of respondents say content creation is now significantly faster, with another 37% reporting it is somewhat faster. This acceleration enables quicker responses to market shifts, customer needs, and competitive pressures, and spans industries. For example, in the retail industry, organizations can produce seasonal promotions within hours. In healthcare settings, providers can generate personalized patient education materials in near real time.

Democratization also extends to collaboration, with 31.8% of respondents saying GenAI has enhanced cross-department collaboration in content creation and 43.5% saying it has done so somewhat. Instant draft generation makes it easier for multiple stakeholders across marketing, product development, and sales to work in parallel, refining and approving content together rather than following a slow, linear handoff process.

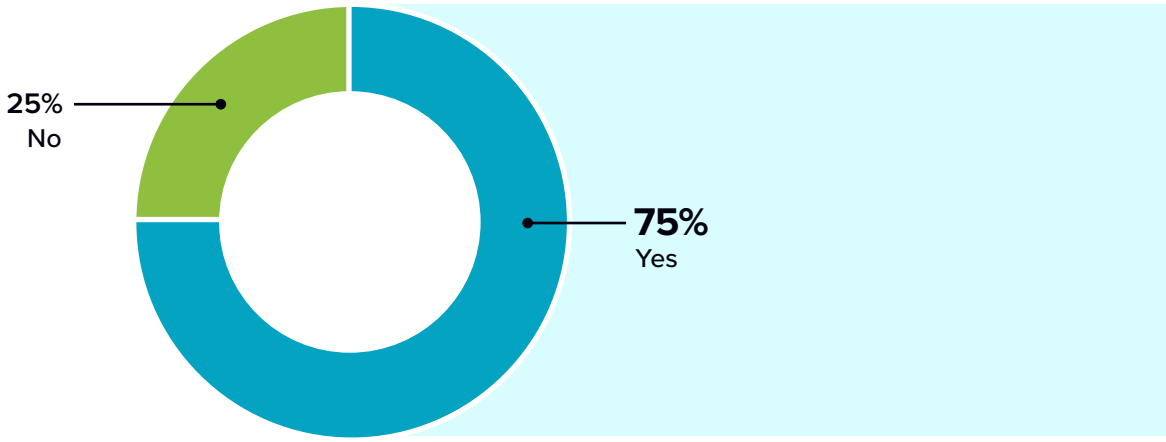
By empowering more employees to create, GenAI increases the potential volume of content. Combined with faster production and richer media formats, this shift can significantly add to data volumes, particularly when organizations store assets for reuse or compliance.

Overall, GenAI's democratization of content is capable of dispersing creative aptitude across the workforce, accelerating production and fostering collaboration. The challenge for enterprises will be to harness this expanded capacity while ensuring that the resulting content and the data it generates deliver measurable business value.

FIGURE 16

Has generative AI allowed employees outside of traditional content or creative roles to now generate content?

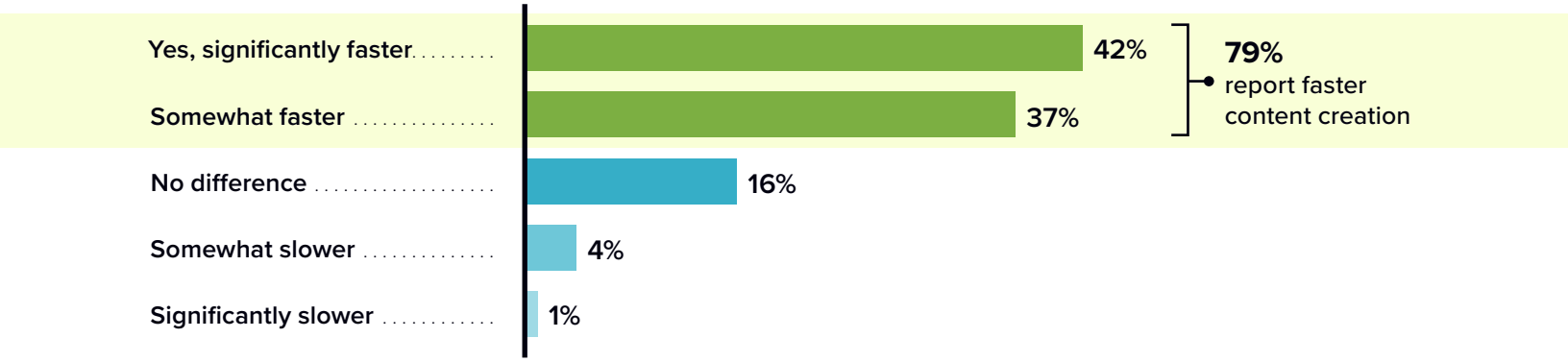
(Percent of respondents)



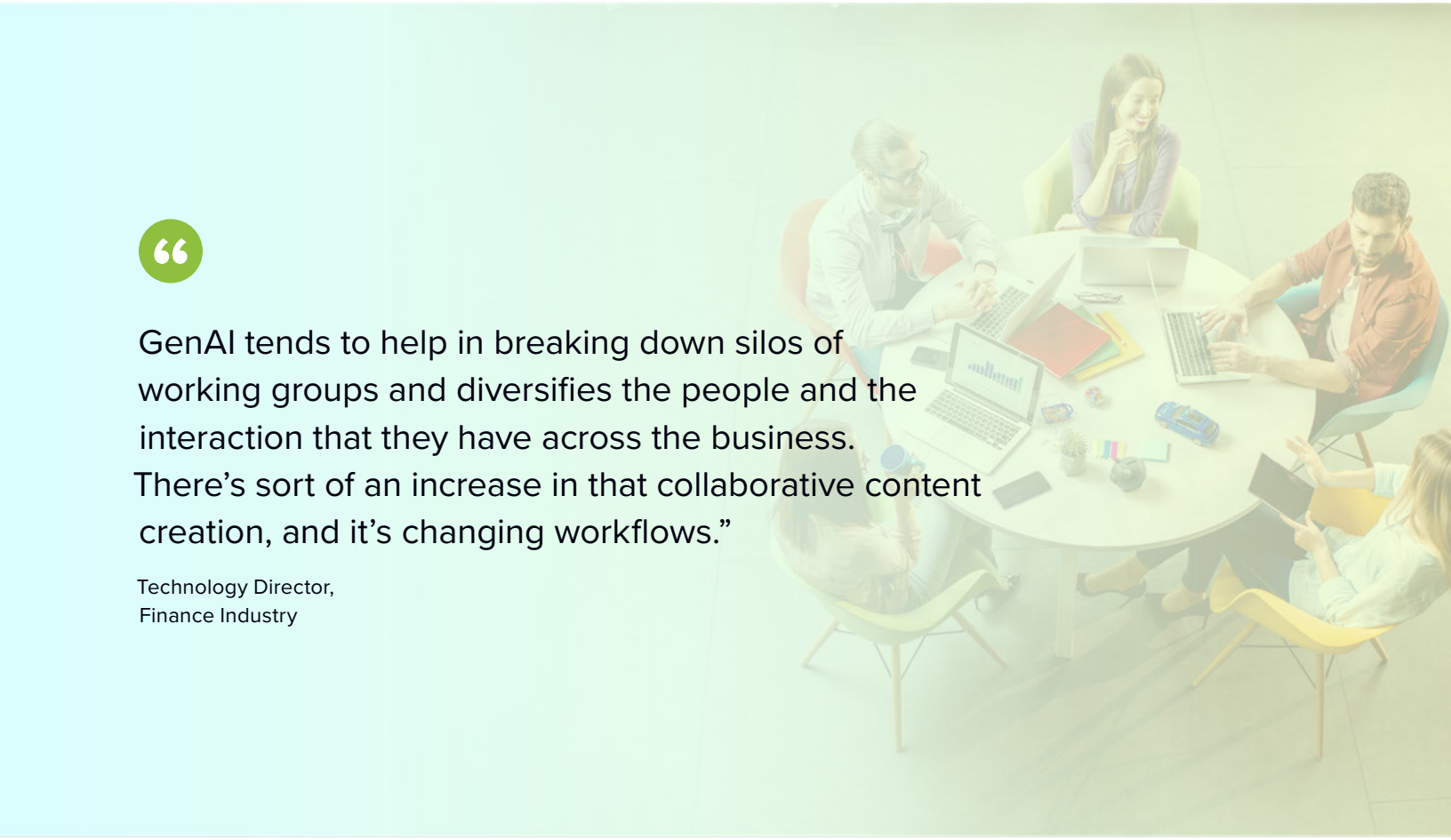
n = 1,034; Source: IDC's *Custom Survey 2025: Seagate GenAI Data Creation Survey*

FIGURE 17

Has generative AI enabled faster creation of content compared to previous manual methods?
(Percent of respondents)



n = 1,034; Source: IDC's *Custom Survey 2025: Seagate GenAI Data Creation Survey*



"GenAI tends to help in breaking down silos of working groups and diversifies the people and the interaction that they have across the business. There's sort of an increase in that collaborative content creation, and it's changing workflows."

Technology Director,
Finance Industry

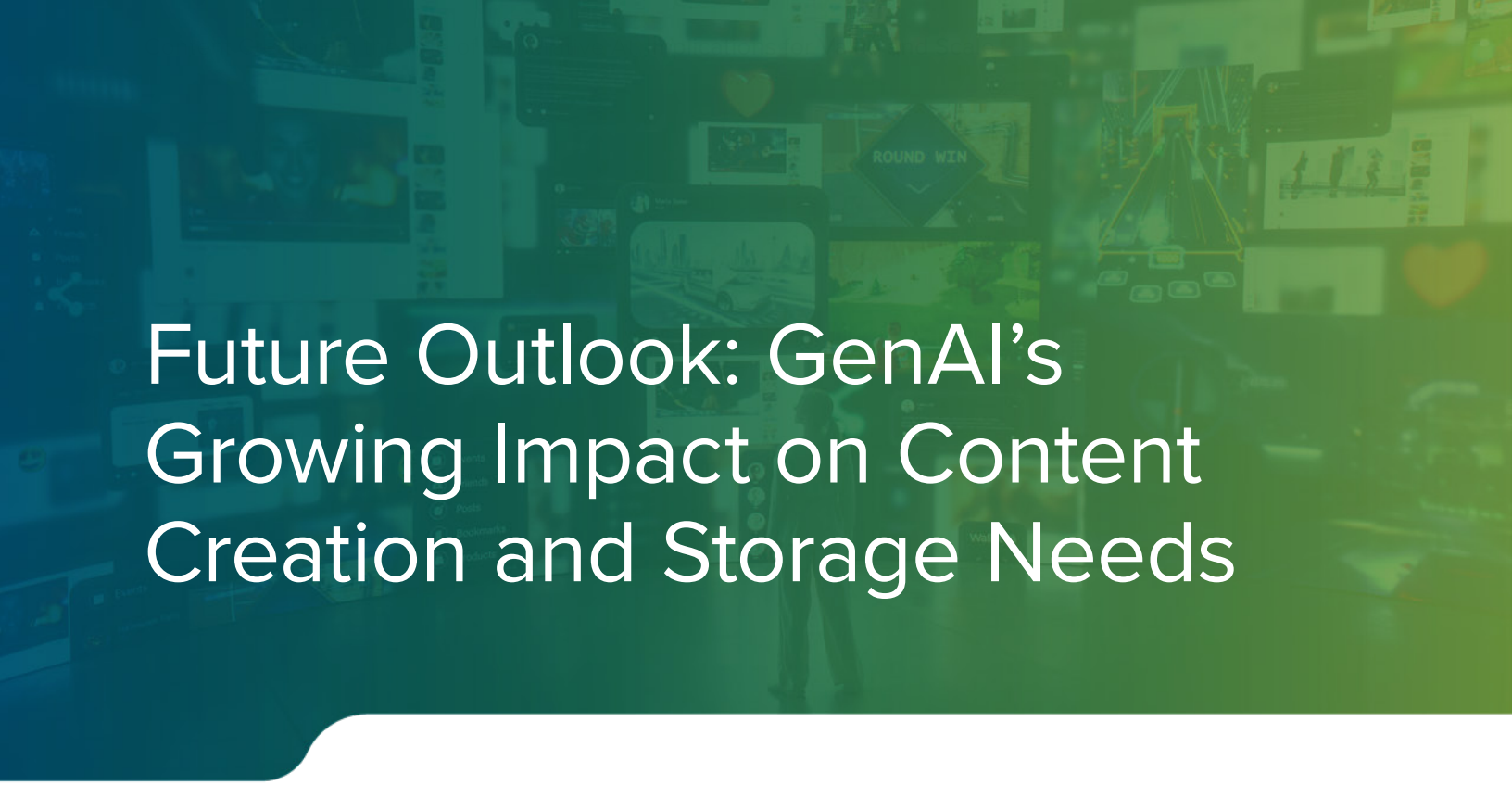
Implications for Value and Scale

Research findings from this study make it clear that GenAI is amplifying the creation and consumption of data inside enterprises.

As content becomes richer, more complex, and more widely replicated, the scale of data growth is accelerating. The key question is no longer whether data volumes will rise but how enterprises can harness that scale to deliver measurable value. Scale is not just an operational byproduct of GenAI adoption; it is a strategic enabler for training better models, uncovering new insights, and delivering differentiated digital experiences. Training models at scale often requires accessing large historical data sets — a scenario where high-capacity HDD bandwidth and efficient TCO can offer the optimal performance needed for ongoing model development.

The value of data in this new environment is multifaceted. For example, organizations are beginning to reuse archived design files, documents, or recorded conversations as training material for domain-specific AI models. Customer service transcripts can be repurposed into knowledge bases that power more accurate chatbots, while video surveillance or sensor data can feed computer vision models to improve safety, efficiency, or customer engagement. These examples highlight that what was once considered “cold” or low-priority data may now hold new value when scaled and applied to generative AI use cases.

Although enterprise adoption of GenAI is still in its early stages, a clear trajectory is emerging: this technology is compelling organizations to store more data, not less. This signals a recognition that every additional byte could hold latent business value, whether in operational efficiency, customer experience, or future innovation. The challenge for leaders is to proactively decide how far to lean into scale to capture that value. Those that modernize storage strategies and embrace data growth as a source of competitive advantage will be in a position to maximize their AI impact, while those that underinvest risk constraining the effectiveness of their models and opportunities for future growth.



Future Outlook: GenAI's Growing Impact on Content Creation and Storage Needs

Generative AI has already begun to reshape the content creation landscape, enabling faster production, broader participation, and more sophisticated outputs across industries.

From marketing and sales materials to creative media, technical documentation, and personalized customer interactions, its ability to accelerate workflows and lower barriers to creation is undeniable. While the scale of impact varies by industry and maturity of adoption, the trajectory points toward generative AI becoming a foundational element of modern content ecosystems and is redefining the value of data.

Looking ahead, the magnitude of this impact could grow significantly. Several accelerators will amplify content creation over the next five years. These include rapid improvements in model capabilities and multimodal AI, which will expand the types of content that can be generated and the speed at which it can be produced. Increasing the integration of GenAI tools into enterprise software platforms will make AI-assisted creation a default feature of many business workflows. For consumers and prosumers, continued embedding of AI into social media, video editing, and creative tools will fuel an ongoing surge in user-generated content. As AI models become more specialized, they will be able to produce higher-quality, domain-specific outputs with minimal prompting, further reducing the skill threshold necessary to generate professional-grade material.

These trends carry clear implications for storage demand. More frequent creation of high-resolution images, videos, and other media-rich outputs increases the data footprint of content, even before factoring in the potential for longer retention of AI-generated material. Storage providers, particularly those serving hyperscale environments, are likely to see incremental demand from consumer-driven and enterprise-driven AI workloads. Over time, if organizations embrace retention of AI outputs for reuse, training, and analysis, this could put sustained upward pressure on storage capacity needs.

In the near term, the impact of generative AI on content creation will continue to accelerate, and the evidence suggests that storage needs will follow. While some content will remain ephemeral and not all organizations will expand retention, the scale, richness, and accessibility of AI-assisted creation point toward a surging growth in digital content ecosystems. As adoption deepens and output volumes grow, generative AI is likely to become a significant driver of content generation and the storage infrastructure required to support it, particularly for providers in a position to serve high-capacity, media-rich workloads. Most importantly, the scale of content creation is a catalyst for unlocking new business value from data that enterprises might have once overlooked.

Recommendations

Generative AI is emerging as a catalyst for unprecedented growth in digital content creation and value, with ripple effects that extend across the entire data ecosystem.

IDC recommends that organizational leaders and IT decision-makers consider the following:

Align GenAI adoption with future-proof, scalable storage architectures:

Plan now for the increase in high-volume, high-resolution content by adopting tiered storage platforms, designed with a high-capacity HDD foundation, that can efficiently scale for long-term demand. HDDs will continue to provide the vast majority of exabyte storage in enterprises over the long term due to their superior value proposition for efficiency and long-term data retention. As AI workloads continue to expand in size and complexity, most enterprise environments will utilize tiered storage architectures that combine SSDs and HDDs. SSDs will continue to be complementary to HDDs.

Redefine what "performance" means in storage architecture:

GenAI workloads will demand more frequent and heavier write operations as the value and volume of data sets grow. HDDs generally have better write endurance compared to QLC NAND, an important aspect of AI workloads today and into the future. It will be important to consider this during the data preparation or synthetic data generation phases of the AI life cycle due to their highly write-intensive nature. The traditional definition of performance storage media should no longer be limited to device latency but extend to critical performance criteria, such as throughput, write durability, long-term data retention, and efficiency at fleet scale.

Enhance governance and retention planning:

Anticipate that some AI-generated content will have long-term value for analysis, model training, or compliance, and define clear life-cycle policies to determine what is worth retaining and for how long.

Leverage hybrid and cloud storage models:

Combine on-premises HDD storage for predictable workloads with cloud HDD capacity for elastic scaling during content surges, ensuring resilience and economic efficiency.

Prepare teams for AI-driven content workflows:

Invest in training so that employees understand not only how to create and manage GenAI content but also how storage strategies impact accessibility, performance, and cost over time.

By anticipating the storage implications of AI-driven content creation today, organizations can ensure they have the necessary infrastructure in place to support tomorrow's workloads, positioning themselves to fully capitalize on the next wave of digital content growth and extract more value from their data.

Methodology

The viewpoints, analysis, and recommendations that this white paper presents are based on IDC's market research into the impact of generative AI on content creation and storage needs.

Part of the research included a survey of over 1,000 business and IT decision-makers worldwide from large and midsize organizations across industries. In addition to the survey, IDC conducted several in-depth interviews with enterprise end users of GenAI tools and relied on existing IDC syndicated research, such as IDC's Global DataSphere and Global StorageSphere, about worldwide data generation and storage dynamics.

About the IDC Analysts

Adam Wright
Research Manager,
Global DataSphere and Global StorageSphere, IDC

Adam Wright is a research manager for IDC's Global DataSphere and Global StorageSphere research programs. Adam is responsible for the Global DataSphere forecast, which measures the total volume of data created each year across the world, and for the Global StorageSphere forecast, which is a measure of the installed base of storage capacity worldwide, and the amount of data stored in any given year. Additionally, Adam leads insightful research that explores key trends, use cases, technologies, and other factors shaping both the Global DataSphere and StorageSphere.

More about Adam Wright

Greg Ireland
Senior Director,
Consumer Markets and GenAI for Content Creators and Consumers, IDC

Greg Ireland is a senior director for the Consumer Markets programs at IDC. In this role, he manages IDC's Consumer Market Trends, Consumer Market Model, and GenAI for Content Creators and Consumers research programs. He focuses on consumer adoption of and engagement with digital technologies, services, and applications that transform consumer experiences, business models, and market opportunities. Greg leads IDC's coverage of consumer GenAI, and he also has expertise in and provides in-depth analysis on the ways in which digital video content is distributed, consumed and monetized across traditional pay TV, over-the-top (OTT), and social media services and platforms.

More about Greg Ireland

Edward Burns

Research Director, Hard Disk Drive and Storage Technologies, IDC

Edward Burns is IDC's research director for hard disk drives, responsible for providing insight and market data for the worldwide HDD market. Mr. Burns' research coverage includes all market applications for hard disk drives as well as assessing demand for key HDD components. Additionally, Mr. Burns' research includes analyzing hard disk drive technology trends, technology trends for competing storage technologies, and the market opportunity for all forms of storage technologies across all markets worldwide.

More about Edward Burns

Message from the Sponsor

Seagate is a leader in mass-capacity data storage.

We have delivered more than four and a half billion terabytes of capacity over the past four decades. We make storage that scales, bringing trust and integrity to innovations that depend on data. In an era of unprecedented creation, Seagate stores infinite potential. To learn more about how Seagate leads storage innovation, visit www.seagate.com and our blog, or follow us on X, Facebook, LinkedIn, and YouTube.

Learn more about how Seagate leads storage innovation

Visit our blog

IDC Custom Solutions

IDC Custom Solutions produced this publication. The opinion, analysis, and research results presented herein are drawn from more detailed research and analysis that IDC independently conducted and published, unless specific vendor sponsorship is noted. IDC Custom Solutions makes IDC content available in a wide range of formats for distribution by various companies. This IDC material is licensed for external use and in no way does the use or publication of IDC research indicate IDC's endorsement of the sponsor's or licensee's products or strategies.

IDC Research, Inc.
140 Kendrick Street, Building B, Needham, MA 02494, USA
T +1 508 872 8200

idc.com @idc @idc

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications, and consumer technology markets. With more than 1,300 analysts worldwide, IDC offers global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries. IDC's analysis and insight helps IT professionals, business executives, and the investment community to make fact-based technology decisions and to achieve their key business objectives.