



There's Nothing Artificial About AI for Broadcasters:

NEW SURVEY SHOWS RAPID UPTAKE

The media and entertainment industry is in the early stages of adopting artificial intelligence for applications ranging from speech-to-text caption generation to automating metadata tagging chores. Some early AI adopters even have begun to monetize AI technology by making it easier and faster to find, retrieve, and reuse archived video clips.

SPONSORED BY

Quantum

IN PARTNERSHIP WITH

TVTechnology
Broadcast • Production • Broadband • Satellite • Mobile

EXECUTIVE SUMMARY

In today's popular media, there has been an explosion of coverage about artificial intelligence (AI). Stories range from the enthusiastic—describing a fantastical utopia of intelligent assistants that answer humankind's every request—to strident editorials calling for rejection of any new technology that threatens to replace human jobs. Hollywood has contributed to the hype surrounding intelligent machines with a variety of tales about many possible futures with powerful AI systems, both good and evil. Of course, reality lies somewhere between AI nirvana and mass unemployment and poverty. As research progresses, it becomes feasible to apply the technology to more and more applications, but the science is a long way from creating self-aware AI über-beings.

AI is a broad field that covers a range of loosely related technologies. Generally, AI is defined as the ability to perform tasks that require human-like intelligence, including image and speech recognition, logical deduction, and the ability to make decisions based on past experience. Rather than simply executing a fixed set of procedures that have been hard coded into software, AI-based systems are able to “learn” about tasks and data sets, and adapt their behavior to optimize outcomes and deal with changing inputs. Machine learning and cognitive analytics are other fields that are closely related to AI; for the purposes of this report, these terms will be considered synonyms.

To better understand the degree to which the media and entertainment (M&E) industry has adopted this technology, TV Technology magazine conducted a survey of 300 M&E business and technology professionals. The survey was underwritten by Quantum, a leading expert in media workflow storage, archive, and data protection based in San Jose, California. Quantum had input into the survey questionnaire, but TV Technology had final approval on the survey. This white paper provides an analysis of the responses to help provide an accurate picture of how widely AI technologies have been tested and put to use in actual production environments.

Two-thirds of the organizations that have tested AI technologies have adopted them for automating at least one significant step in their workflow, such as captioning, metadata tagging, or social media clip generation. Media companies that have large content libraries and are “metadata rich” are particularly enthusiastic adopters.

Contrary to much of the current industry hype, most of the heavy lifting for content analysis and storage today is being done using on-site and off-site private cloud resources controlled by content owners. Public cloud utilities have a presence in this market, but they are only being used by a minority.

Survey respondents project future growth to be robust, both in terms of the amount of content being created as well as the overall size of content archives.

APPLICATIONS FOR AI

M&E companies, like most other businesses that need to produce a profit, are taking a hard-nosed approach to AI. For any new process to be adopted, it must produce measurable value and be able to withstand the stresses of a relentless, demanding workflow. Many of the surveyed companies have moved beyond initial trial periods and are using AI as part of their production system. Their attitudes, as reflected in the results of this survey, indicate that AI is not a solution in search of a problem. Rather, AI is a tool that can be deployed successfully at scale today.

So which applications in the M&E industry are being implemented with the help of AI? The answers fall into a few broad categories.

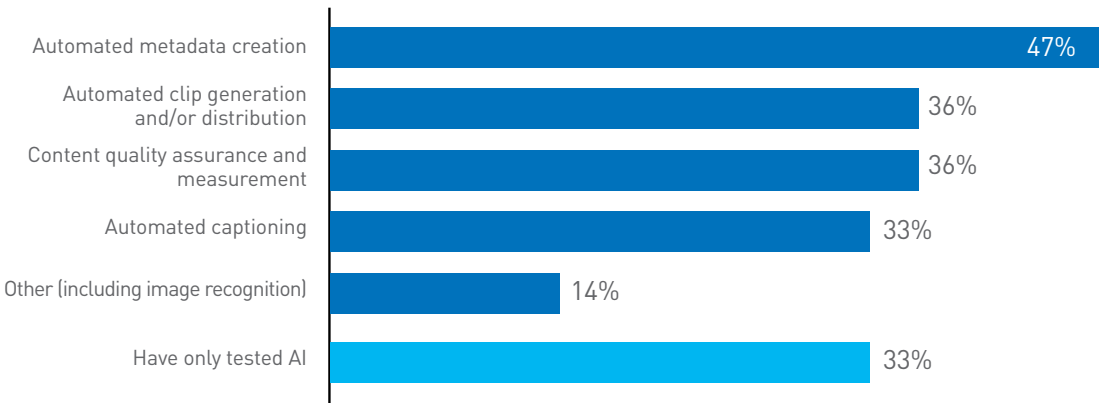
- Most prevalent is automated metadata creation, where clips, streams, and files are examined for audio or video characteristics that match a variety of criteria for attaching metadata tags or labels to the content for the purpose of cataloging and later retrieval. This process saves time and alleviates a great deal of the drudgery that is required to tag sports and news content.
- Another popular AI application within the M&E industry is automated clip generation and distribution, where content files and/or live

streams are analyzed for characteristics that may be popular with viewers or that can drive lots of attention on social media. Once identified, these clips can be transferred to the specific file formats required for each social media platform and automatically posted for viewing, along with relevant comments and tags.

- Content quality assurance and measurement also can be enhanced with AI techniques. For example, verifying that captions are present is a relatively simple task for an analysis tool, but verifying that the captions correspond to the dialogue or narration, and that they are properly synchronized, requires a much higher level of cognition.
- For that matter, fully automated captioning is another significant goal for AI technology. Current systems are not yet 100% accurate, but progress towards that goal continues to be made. Machine-assisted captioning in multiple languages is another viable technology, with particular application to captioning archived materials. For example, this process can be used to apply captions to stored materials that may have been broadcast before current regulations were in place. Once an archived clip has been captioned, it can be rebroadcast more easily today. In addition, the captions themselves provide a rich set of data for searching within content libraries.

Figure 1

How organizations with AI experience are currently using the technology.



According to the survey results, AI is starting to have a meaningful impact on media production practices. Already, 21% of the survey respondents have had experience with AI in their own operations. Of these, almost half use it for automated metadata creation, as shown in figure 1. One-third use AI for one of the other three applications previously mentioned, including clip distribution, quality measurement, or automated captioning, and some organizations use it for all three purposes. As shown in figure 1, one-third of the companies that have AI experience have so far only experimented with the technology, but are not yet using it, which is to be expected for a relatively new field. It will be interesting to see how rapidly other organizations try, and then possibly adopt, this technology as it matures and becomes more mainstream within the M&E industry.

LARGE CONTENT LIBRARIES

More than three-quarters of media companies with large content libraries (those with more than 20,000 hours) currently use AI for automated metadata creation. This should not be surprising, as these organizations are also enthusiastic users of other workflow automation tools.

The survey results show that large library owners in general embrace efficient systems for managing their content, as figure 2 shows. The major commonalities among the organizations surveyed that have large content libraries include:

- Three-quarters have a process for regularly or automatically migrating content from older/obsolete storage technologies. This essential activity helps to prevent content from becoming lost or irretrievable due to the inability to read media that are not supported by current devices (such as old tape formats).
- Almost two-thirds currently use AI for automated clip generation and/or distribution. This process helps to simplify and speed up the workflows associated with publishing program excerpts across a diverse collection of social media platforms, as well as on websites controlled by the organization and related collaborators.
- Three-fifths of those with large archives use media asset management (MAM) systems to perform tasks such as automated archiving, cross-platform file management, and metadata management. These systems help ensure that content is available when it is needed, stored economically, and secured against unauthorized access.

Figure 2

Organizations with large content libraries use advanced tools, including AI technologies.



- For 68% of these organizations, more than half of their content is properly tagged with metadata and could be retrieved within 10 minutes. To better understand what this means, think about the steps involved. First, a search must be initiated using some key words or other tags that are related to the desired content. Then, candidate clips must be located within the various tiers of the storage system and gathered into a selection list. Proxies (if available) may need to be cued for viewing. Clips that are chosen then need to be retrieved and made ready for playback. Doing all of this in less than 10 minutes for a library containing more than 20,000 hours of video is pretty amazing and certainly requires an advanced level of automation and a high degree of consistency in tagging and labeling throughout—nearly perfect applications for AI.

CONTENT STORAGE PREFERENCES

In order to perform content analysis, AI systems need to be able to access the files and streams that are to be processed. There are two locations that need to be considered: the physical location where the content is stored, and the place where the analysis algorithms can utilize processing resources to perform their operations.

These locations don't necessarily need to be the same. With high-speed network connections and process virtualization, location is not as important today as it once was. Nevertheless, as can be seen in figures 3A and 3B, these organizations tend to store and analyze their files in similar locations. Across all major user categories, the most popular location for cloud storage (unstructured object storage) is on-site private storage. Almost half of the respondents use this kind of storage. Off-site private cloud storage is less prevalent, at 18%. The remaining third of the survey respondents reported using public cloud storage.

For AI-driven analysis, on-site private resources are also most popular, with more than half of the respondents citing this as their preferred location. One-quarter reported using off-site private cloud storage.

It's interesting to note that the public cloud (i.e., cloud storage and processing services available from providers such as Amazon Web Services (AWS), Microsoft, and Google) is used by only a minority of respondents, for both storage and analysis. This may be somewhat surprising, as much of the published material that describes AI-like technologies talks about using public clouds for performing the required analysis. There are several factors that could play into the apparent preference for private clouds (both on-site and off-site), including:

Figure 3A

Preferred locations of unstructured object storage clouds.

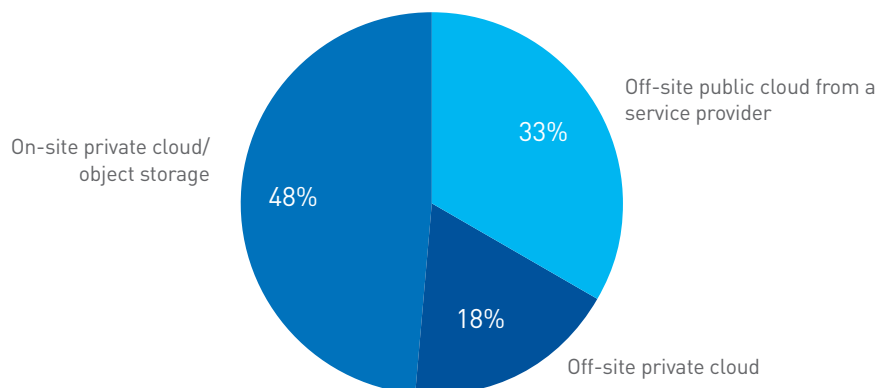
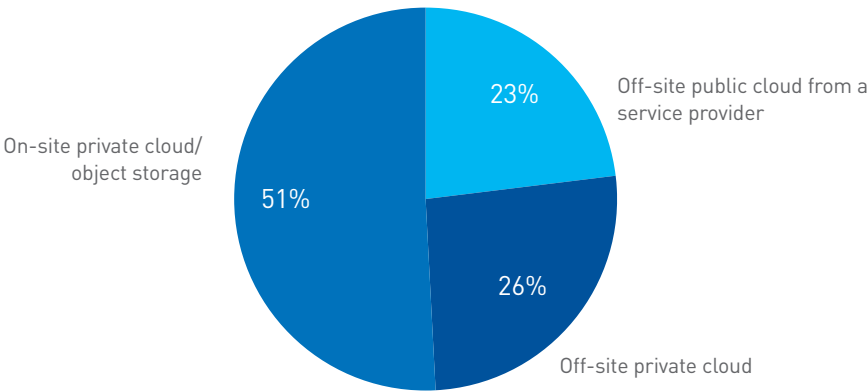


Figure 3B

Preferred locations for performing AI content analysis.



- Cost differences between private cloud and public cloud services, particularly when the expense of high-bandwidth Internet access is considered for sending and retrieving files.
- Security and privacy may be considerations in deciding where files are located, and particularly if content ownership agreements contain clauses about where specific content files can be stored.

GROWTH EXPECTATIONS

According to survey respondents, the pace of content creation is continuing to accelerate, which is creating further pressure to automate content workflows using AI, among other technologies. Survey results clearly indicate that the amount of content being managed will continue to grow. Most tellingly, only 2% of respondents think their

archives will decrease over the next three years and only 3% feel their organization will create less content over the next three years. In contrast, as shown in figure 4A, three-quarters of respondents expect the amount of content they produce to grow by more than 10% in three years. Figure 4B shows that one-third of respondents think their archives will increase by more than 50% in the next three years (as measured in hours of content).

Neither one of these results should be particularly surprising, although the sheer scale of anticipated growth may be greater than expected. Storage costs have continued to decrease, and the number of potential avenues for content reuse and monetization has increased, driving more organizations to create and maintain larger content archives than they did previously. The costs for producing each hour of content have also followed

Figure 4A

A substantial majority of organizations expect to be producing more content over the next three years.

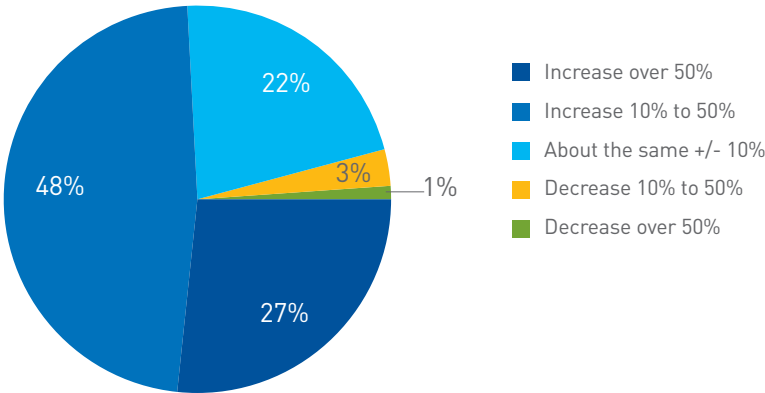
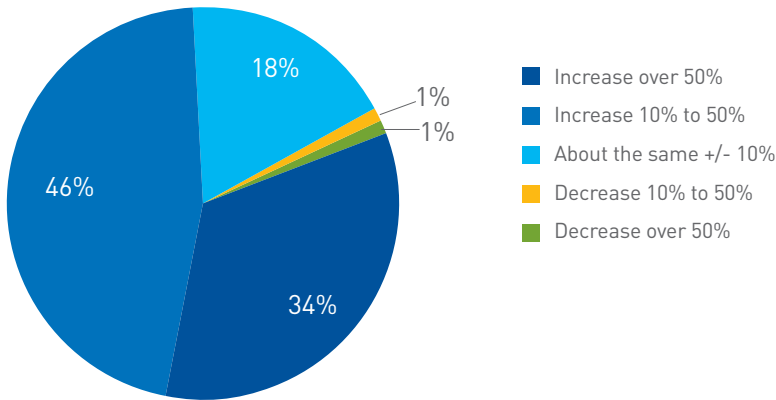


Figure 4B **An even larger majority expects archives to grow over the next three years.**



a downward trend, enabling more content to be created without increasing production budgets.

AI USAGE BY TYPE OF ORGANIZATION

Respondents from a wide range of different categories currently use, or have at least tested, an AI technology. As figure 5 illustrates, the percentage varies from 10% to nearly 45% for each of the major categories, except one. This is fairly good news for advocates of the technology, as it indicates that there are many potential future users that still need a proper introduction to the AI solutions available.

Only one of the surveyed user categories, broadcast TV stations or groups, did not indicate any experience with AI. At first, this response seems surprising. Why wouldn't a local television broadcaster want to simplify its process for generating metadata for the content it produces? A deeper look reveals that broadcasters overall are pretty satisfied with the amount of metadata they're creating for new content. In fact, 20% said they have enough metadata for their entire library, and 53% indicated they have enough metadata for newly created content, but not for older material. In another part of the survey, only 10% of the organizations in this user category indicated that they're currently using fully

Figure 5 **Respondents that currently use or have used AI technology (by category).**

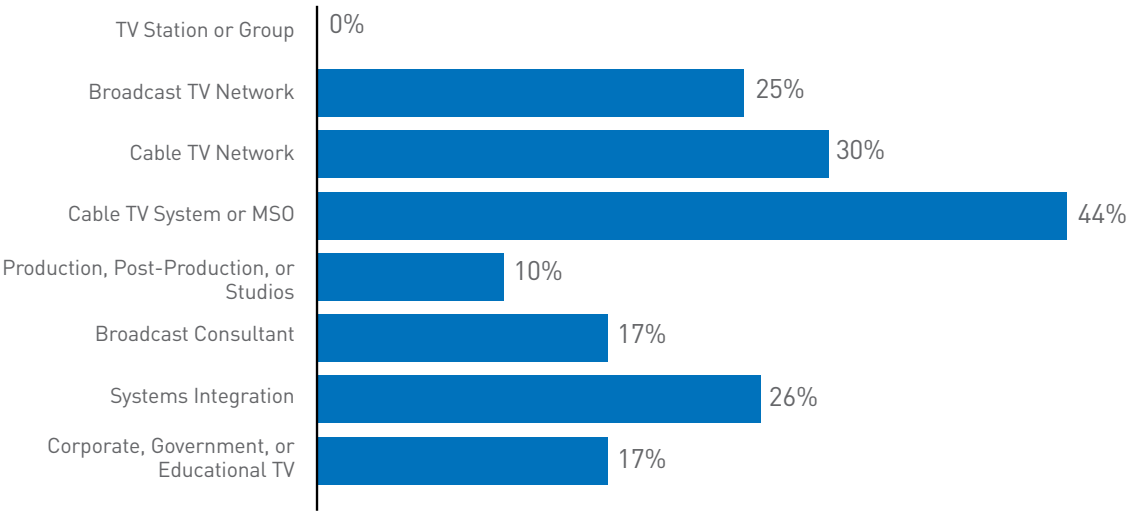
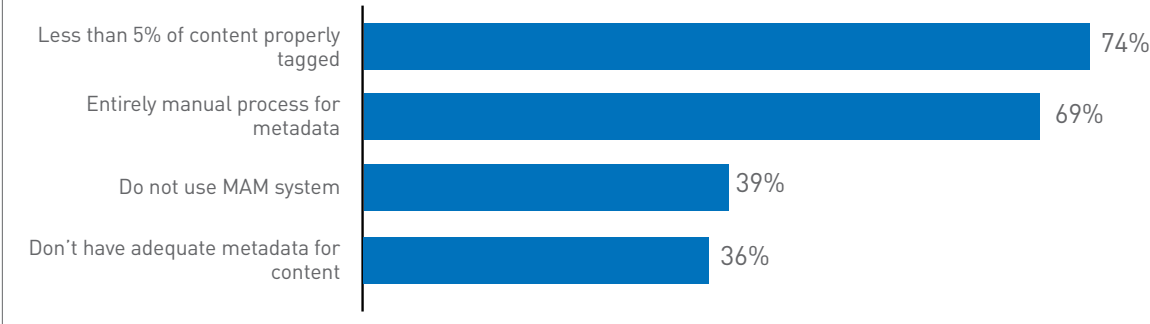


Figure 6 **Common deficiencies connected with limited metadata.**



automated metadata creation systems. So this group is probably not clamoring for better tools to create metadata.

INADEQUATE METADATA PROCESSES

One area where AI could help M&E companies is in the creation of metadata for new and existing content. There are many benefits to having accurate and complete metadata, including the ability to support rapid retrieval of clips from archives to reuse for breaking news stories or emerging trends on social media. In most cases, organizations that can create and post accurate, informative clips in a timely manner will end up garnering a larger share of views—simply by being first. With the right content and the right message, broadcasters and other media companies can jump-start the process of driving viewers to their websites to achieve greater levels of engagement, awareness, and viewership.

Unfortunately, many organizations do not have adequate metadata processes in place. While they may have relevant content sitting in their archives, the inability to find and retrieve relevant content quickly could mean missing out on a rapidly trending topic in social media. Without the right metadata, the process of locating and publishing the content becomes a much more difficult and time-consuming task. The same can be said for traditional broadcasting, because more effort is required to locate old news clips, sports stories, and possibly even footage that could be used for new entertainment productions.

To help illustrate the current limitations of metadata, figure 6 graphs a few data points extracted from the overall survey respondents:

- Almost three-quarters of respondents feel they do not have adequate metadata in their content library. Apparently, the disadvantages of not having well-tagged content files are understood, but perhaps either the perceived costs or the apparent difficulties of improving metadata coverage are too great for many media organizations.
- More than one-third have less than 5% of their content properly tagged and are not able to retrieve desired content in less than 10 minutes. This is unfortunate, because 10 minutes can mean the difference between being one of the first to report on a new topic and being just another outlet that is joining in after a trend has become widespread.

Storage costs have continued to decrease, and the number of potential avenues for content reuse and monetization has increased—driving more organizations to create and maintain larger content archives than they did previously.

- Almost 40% have an entirely manual process for metadata creation, which means either high recurring costs for content tagging or less thorough content labeling if constrained by budgets.
- More than two-thirds of the organizations surveyed do not use a MAM system, which means that the process of controlling how content files are stored, archived, and secured is a problem for human staffers instead of automated systems.

ORGANIZATIONS WITH ADEQUATE METADATA

A quarter of the survey respondents reported having adequate amounts of metadata for their content libraries. This group of respondents was representative of the general survey population and included organizations from all of the different user categories. Some of these organizations create large amounts of content each year, and others only a small amount. Some have a great deal of content in their archives and others do not. Some use entirely manual processes for metadata creation while others use automation.

Taken as a group, metadata-rich organizations are more than three times as likely to have used AI technologies as those organizations that feel that

they do not have adequate metadata—43% versus 14%. Those that have tried AI are more likely to have gone beyond the realm of testing the technology and started to use it, as figure 7 shows. More than half (56%) of these organizations use AI for automated metadata creation, and two-fifths use it for automated clip generation and distribution.

Of course, metadata is only useful when it is being employed to achieve a business goal. Organizations in this group are able to exploit their archives quickly and efficiently. Almost two-thirds of this group believe they could retrieve a desired clip within 10 minutes from over half of their archive. More than 90% of this group could locate, retrieve, and ingest a four-year-old clip from their archive within one hour. The resulting productivity improvements and faster reaction times of these metadata-heavy organizations offer significant advantages in the race to keep up with the fast-paced world of social media.

USING ARCHIVED MATERIALS

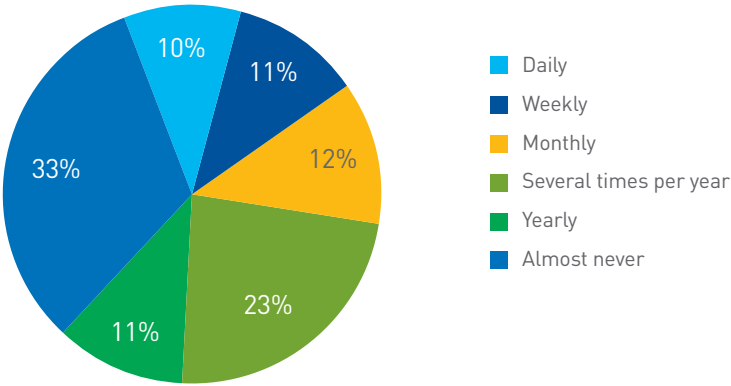
A significant number of organizations routinely use content that has been in their archives for years, as figure 8 shows. One-third of respondents reported using three-year-old content at least once a month, with more than half of that group using three-year-old archived footage weekly or daily. Clearly, these organizations are getting value from their archives through regular use.

Figure 7

Organizations that have tried AI technologies are more likely to incorporate them in their workflows.



Figure 8 Frequency of usage of archived content materials over three years old.



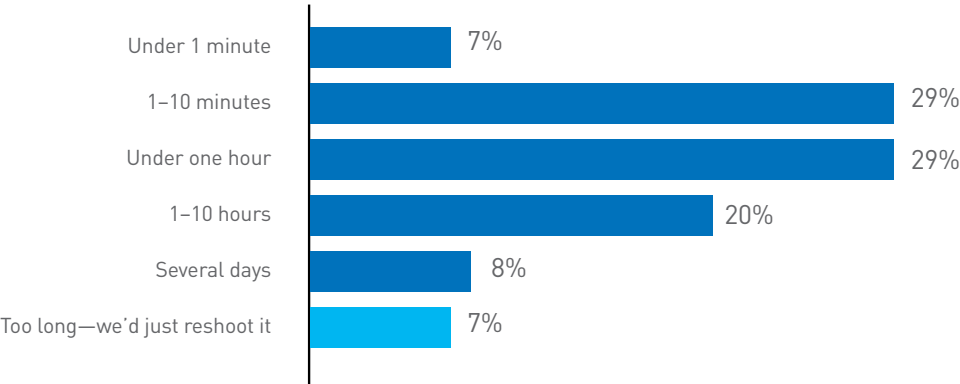
Another third of organizations reported using three-year-old content less than once per year. This latter statistic is surprising in light of the fact that, according to the survey, the average cost of long-term storage is \$96 per terabyte per year, with each terabyte holding, on average, 11 hours of content. This adds up to an annual expense of \$175,000 for a 20,000-hour content library, or \$8.75 for each hour of content every year. This appears to be a considerable recurring cost for an asset that's used infrequently, and particularly for those organizations that use older content "almost never."

ARCHIVE RETRIEVAL TIMES

The time required to retrieve older content from an archive varies greatly from one organization to the next. When asked about the amount of time required to locate, retrieve, and ingest an important four-year-old clip from their content library, answers from respondents covered a full gamut of time frames, as figure 9 shows.

One surprising response was from those organizations that can retrieve a clip from their archive in less than a minute. That is amazingly fast, when all of the steps involved are considered—determining the location of the content, possibly mounting the right media container (tape cartridge, etc.), reading the media, and delivering it to the requesting device.

Figure 9 Average time to locate, retrieve, and ingest a four-year-old clip from the archives of organizations surveyed.



Such speed is not solely an attribute of small operations. Some respondents have speedy archives with more than 50,000 hours of content.

Another group of respondents have very slow retrieval times of up to several days. This could be due to the amount of time it takes to search for the content or due to the amount of time required to access the content in the archive. One potential explanation for this amount of time could be that the archives are stored off-site (possibly in a secure location) and may require a courier dispatch to access the physical media.

Finally, a few respondents (about 7%) indicated that it would take so long to retrieve a four-year-old clip that it would make more sense to shoot the material again. This brings into question the whole idea of maintaining an archive. Does it make sense to spend money on something that is so time-consuming to access that it is rarely, if ever, used?

CONCLUSION

AI has begun to make significant inroads into the M&E industry. As the results of this survey clearly show, AI is not just a set of tools looking for a problem to solve—it's a technology that's currently being used across most user categories in the M&E industry. Of those organizations that have tried the technology, more than two-thirds have used it for applications such as automated metadata creation, automated captioning, quality assurance, and automated clip generation and distribution.

These real-world solutions are making a difference, allowing content producers to reduce the amount of resources required for repetitive tasks. Organizations that have adopted AI are more likely to have adequate metadata throughout their archives (allowing them to know what they have in their back catalog), and they are better able to retrieve clips quickly and derive value from these archives. Investing in archives makes sense, and investing in properly organized and labeled archives makes even more sense.

ABOUT QUANTUM

Quantum is a leading expert in scale-out tiered storage, archive and data protection. The company's StorNext® platform powers modern high-performance workflows, enabling seamless, real-time collaboration and keeping content readily accessible for future use and re-monetization. More than 100,000 customers have trusted Quantum to address their most demanding content workflow needs, including top studios, major broadcasters and cutting-edge content creators. With Quantum, customers have the end-to-end storage platform they need to manage assets from ingest through finishing and into delivery and long-term preservation. See how at www.stornext.com.

©2017 Quantum Corporation. All rights reserved.

Quantum®

www.quantum.com
800-677-6268

ST02194A-v01 Dec 2017