More than ever

**content is king**

AI enables monetization of media archives

By: Thomas Burns, CTO, Media and Entertainment, Dell EMC

INCLUDING: 5 tips to accelerate your AI strategy from GrayMeta CTO John Motz
Comments like this are all too common in the industry as media organizations have woken up to find their organically grown, sprawling content archives missing a key bit of information: metadata.

Television broadcasters and other owners of large content libraries are facing this problem due to the sheer volume of media assets locked up on data tape, with incomplete or idiosyncratic information about exactly what has been stored. Without complete and accurate metadata, it’s difficult to make decisions about the worth of a given media asset, and that makes content libraries difficult to monetize.

How many times have you heard of a film restoration archivist (while looking for something else) “finding” a supposedly lost asset in the vaults, and thus making a restoration project more complete, unique, newsworthy, etc.? Many media companies know all too well that the ability to store data versus its value are often misaligned. Companies have long been looking for tools to help drive value from assets that were created and acquired at great expense but that have become “lost” with little hope of being found.
Data-driven content

Long ago (in internet time), we saw the rise of “big data,” and watched it become “analytics,” which has become “deep learning” (a refinement of machine learning), which is itself part of the catchall term called “artificial intelligence.”

But it’s all a variation on the same thing: performing algorithmic queries on deliberately acquired data about customers, products and services in order to yield actionable intelligence.

The media distribution business has embraced analytics. Successful media and entertainment distributors had to develop decision making capabilities allowing them to respond rapidly to constant change, while also integrating rights, digital supply chain, web and social media data. This has helped distributors gain powerful insights about both their media assets and customers, which further informs the types of programming they are willing to invest in.

This use of analytics has opened a window into a critical link between customer preferences, monetization and what is in your media archive. In a recent study by Nielsen of the subscription video-on-demand (SVOD) consumer viewing habits, what’s driving 80 percent of the time spent viewing these services is the back catalog of content acquired by streaming services from television networks and studios. “...Our research shows most of the viewing time is spent with catalog-programming,” said Nielsen’s COO Steve Hasker. In short, the content in media archives is driving the majority of the new and rapidly growing ways consumers are viewing media.
Creating metadata has typically been a manual process, where an informed and knowledgeable person can “tag” (or assign metadata to) a media asset, and armed with an appropriate taxonomy, standardize the description of the asset so that search techniques can be used with confidence.

But this approach doesn’t scale. There aren’t enough qualified people, let alone enough infrastructure to let users watch and tag content interactively. Further, video and audio need to be watched and listened to in real time, and there are literally thousands of years’ worth of material which may be trash—or treasure. No one will know without evaluating the asset.

Every rights holder has had to migrate their data—their companies’ precious capital assets—from an older medium to a newer one, simply to preserve its ability to be read in the future. This is a labor-and time-intensive process, which doesn’t add any intrinsic value to the media assets, but has to be done regardless. What if you could migrate your data once, add value to it as part of the migration, and then never need to migrate to another tape format again?

If you could “automagically” add metadata to your content archive—through a combination of AI techniques that watch and listen to your library content and build up a user-referenceable database of people, places, things, even sentiments—you would have the ability to create a new type of programming, where archival material could serve as context for current-day narratives much more easily than such programming is created today.
As multiple sources of data proliferate within an organization, new AI and machine learning techniques evolve to make cross-correlations visible to departments that previously did not have visibility into this data. This promotes companies’ adoption of a “platform” approach to metadata collection, instead of an application-specific approach, where on-set or production metadata might not be considered useful in the digital supply chain or distribution metadata.

Applications that can analyze and correlate media assets with sources of user data can turn these static media assets into “data capital”—an asset class which will continue generating revenue throughout the life of the asset, much like real estate continuously generates income for its owners. Like real estate, the owners of data capital will have to invest in maintenance in order to keep generating revenue from that asset class. The ability to rerun an updated machine learning algorithm against an existing media asset library, correlated with more recent user data such as social media feeds or sensor data from location-based entertainment, may tease out previously overlooked narratives, or themes that can then inform new uses for the content.

However, the sheer size of media assets can make copying data from a passive archive like a tape library less agile than required for a quick reaction to market or celebrity news. A “data lake,” or globally scalable storage fabric consisting of scale-out NAS and geo-scale object storage systems, allows multiple workloads such as archiving, disaster recovery and collaboration to be executed without requiring multiple silos for each.

Now imagine you can weave an AI appliance into that data lake, which can be trained on your own data capital without having to migrate your data somewhere else for that purpose! This architecture will allow AI algorithms to perform facial recognition, object recognition, audio transcription, and even translation on media assets, in order to identify and track features which are useful to advertisers, piracy watchdogs, organizations tasked with identifying manipulated imagery, and others.

The appearance of toolsets which have moved “up the stack” from raw computer science algorithms, and become software frameworks and applications that run on scalable clustered appliances without having to migrate vast amounts of data has ushered in a new era of revenue generation. This in turn allows content rights-holders to breathe new life into their existing pipeline of content creation, content management and content distribution.

Content and the rise of machine learning
By now most people have heard of artificial intelligence (AI), but they may not know how much they already rely on it. Whether used by a retailer to anticipate customer needs, or a credit card company detecting possible fraud, artificial intelligence enables us to streamline tasks and achieve unimaginable results. But we’re still just scratching the surface of what AI can do.

Here are five tips managers should keep in mind when considering AI adoption for their businesses.

1. **Boost your return on investment by diving in**

Customers often ask, “When should I start?” The answer is now. Even if the improvement in your organization’s ROI turns out lower than expected at 10 or 15 percent, you’re kicking off a process that leads to improved ROI, as you increasingly rely on AI to improve efficiencies.

Often, potential customers make the mistake of trying to understand how the technology operates, instead of focusing on the unique problems of their organization. AI isn’t technology for tech’s sake—it’s a problem-solving solution.
With the right planning and preparation, businesses can immediately see results. If the AI scope is limited to improving a process reliant on task-driven, repetitive work, there’s minimal execution risk, and a high likelihood of success on your first pass.

2. When adopting AI – think big, start small, scale fast
Getting acquainted with AI can be simple if you start small. Take an aspect of your business that requires a lot of repetitive task-driven labor, or focus on solving a specific problem. Build AI to solve that one problem. Your first AI project should take no longer than 90 days to complete.

For example, in recent years some movie studios have been interested in localizing their film trailers. Consider the human hours necessary to search a feature film for moments with an array of emotional cues best matched to 18 different international markets—and this is all before an editor has even begun to cut any trailers together. AI can use facial recognition to analyze footage and isolate scenes with specific emotions—happiness, crying, laughing—then assemble those scenes and quickly spit out a hundred potential trailers.

The editor’s job just went from a monotonous weeks-long labor to a more interesting task: Finding trailers that work best for certain audiences, punching up the winners, and collaborating with marketing professionals on sharpening trailers for specific customer profiles.

Business leaders sometimes encounter pushback about AI from worried staff concerned about losing their jobs to technology. But as the film editor example shows, there’s still plenty for your employees to do. You’re empowering them to serve thoughtful, judgment-based roles instead of tedious, task-driven ones.

3. Talent matters—enlist the right AI pros
The AI expert or solutions firm you hire might not be familiar with your specific industry. But, first and foremost, you primarily need someone who understands AI. Even a supermarket chain would usually be better served by hiring someone with no background in food retail, but lots of experience in AI—rather than someone with a moderate understanding in both areas.

The technology is extremely complex and you can’t cut corners with the talent you hire.

It’s also wise to hire a data scientist who will help your team effectively gather and analyze the information that AI offers (as well as the information feeding the machine learning). Without a data scientist, your team will struggle, whether or not you have the engineers to build the project.

Currently, finding the right people can prove difficult, as AI technology is still relatively new, and outside of larger tech organizations, there’s a shortage of talent. But that should change in the next five years as a new crop of AI pros emerges.

4. Clean up foundational data using machine learning
Foundational (or old) data is the biggest problem that businesses encounter during AI adoption. Data has often been gathered by people over the course of many years, and can be rife with human error and inconsistencies.
Flaws, gaps or irregularities in your foundational data may cause your AI to break down.

While daunting, cleaning up foundational data using machine learning is critical. No AI is involved with this step—machine learning will analyze the quality of your initial data and normalize any discrepancies so that the data will be useable.

When introducing machine learning, start from the ground up. The process can disrupt human workflows that may have been in place for decades. It takes time. There’s no quick fix—rethinking processes can require input and flexibility from multiple stakeholders. You might even have to throw your longstanding data taxonomy out the window. But if you want to attain the full value of AI, be flexible in your approach.

5. Prep now to keep pace with the AI industry
While AI technology has evolved rapidly in recent years, it is now accelerating at an even faster pace. In 2019 alone, expect to see as much progress as we’ve had in the last five years combined. New partnerships and collaborations are forming, and higher-end institutions are encouraging AI and visual engineering fields, so expect a wave of new talent to surface in the coming years.

Computing power is also a key factor in AI’s advancement. For example, consider facial recognition. Five years ago, face-training algorithms required thousands of subject facial images for the base model, as computing power just wasn’t muscular enough to quickly handle and process images. Now, even in a palm-sized iPhone, facial recognition is a standard feature, and takes place rapidly.

Innovations are emerging faster than ever, and AI’s pace and skills are fed by adoption. The more we use artificial intelligence, the more it improves. Recent AI and machine learning developments have been exciting, but there’s much more to come. Hold on tight — this ride is just getting started.