



SPEECHMATICS

SPEECHMATICS
SMART GUIDE

HOW TO

Amplify your Digital & Media Asset Management using Voice Technology

5 BENEFITS OF USING VOICE TECHNOLOGY TO
IMPROVE YOUR DIGITAL & MEDIA ASSET MANAGEMENT

WWW.SPEECHMATICS.COM

Table of Contents

Introduction 3

 What is the difference between digital & media asset management? 4

 The growth of digital & media asset management 6

 Why is digital & media asset management important to your organization? 8

 The power of voice technology 10

Getting started 12

 Audio and video assets vs image assets 12

At a glance 13

 1. Speed and accuracy 14

 2. Time and cost optimization 15

 3. Content unification 16

 4. Enriched metadata 17

 5. Effective archiving 18

Conclusion 19

Introduction

This smart guide explains the 5 benefits of using voice technology to improve your digital & media asset management.

**THE
DIGITAL ASSET
MANAGEMENT MARKET
IS PROJECTED TO GROW
FROM USD 2.44 BILLION IN
2017 TO USD 5.66 BILLION
BY 2022, AT AN EXPECTED
CAGR OF 18.3%.
MARKETSAND
MARKETS**

What is the difference between digital & media asset management?

First things first, we need to establish the difference between digital asset management (DAM) and media asset management (MAM) and explain why we have grouped them in this smart guide.

Digital asset management (DAM)

Digital asset management (DAM) is a business process for organizing, storing and retrieving rich media and managing digital rights and permissions. Rich media assets include photos, music, videos, animations, podcasts and other multimedia content.

"DAM involves the creation of an archive, the development of an infrastructure to preserve and manage digital assets and a search functionality that allows end users to identify, locate and retrieve an asset. At its simplest, a DAM is a set of database records. Each database record contains metadata

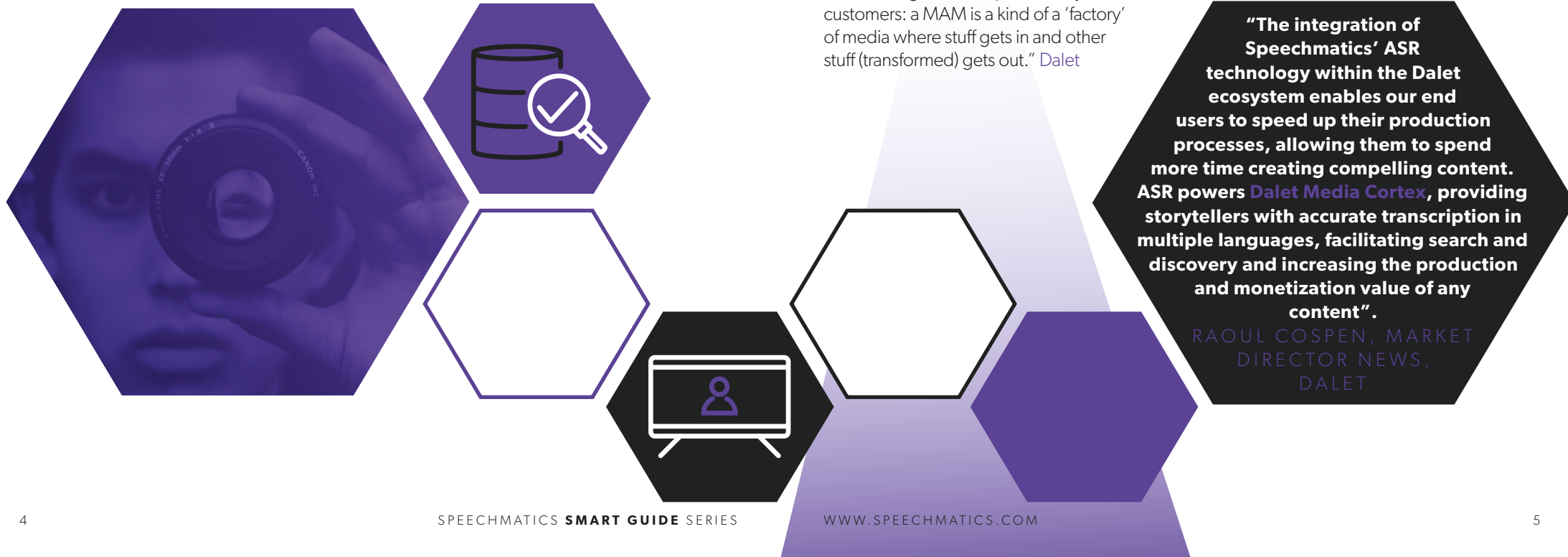
explaining the name of the file, its format and information about its content and usage. Digital asset management software can be used to create and manage the database and help the company to store rich media in a cost-effective manner." [Tech Target](#)

Media asset management (MAM)

"Organizations want to manage their media assets from ingest to distribution. An organization has more and more media to ingest and more and more media to distribute and is looking for a solution. A good example used by our customers: a MAM is a kind of a 'factory' of media where stuff gets in and other stuff (transformed) gets out." [Dalet](#)

Media content management

While we recognize digital asset management and media asset management are different types of media content management, when it comes to voice technology, the same benefits can be applied to both processes. It is also important to note at this stage that most businesses will do both digital & media asset management as part of their media content management and processes.



The growth of digital & media asset management

Digital & media asset management solutions have become increasingly popular for all businesses due to the increase in the volume of digital assets that are created daily. With a monumental rise of digital data that is created from so many existing and new channels, it's no surprise that organizations are having to seek out new tools and solutions to manage existing assets and new ones that are continually produced.

According to a [Cisco 2019 report](#), video content captures four-fifths of global internet traffic. Cisco stated that the dramatic increase is not only driven by the increased popularity of over-the-top (OTT) video streaming services, but also by the number of people expected to be connected by 2019. Now, over half the globe (almost 4.57 billion people) have access to the Internet; however, the number of devices able to access the web is up to three times as high as the global population.

These statistics of growing video adoption highlight the growing need for advanced solutions for broadcasters and media facilities to improve operational efficiency by streamlining their workflows.

Companies including both large and small to medium-sized enterprises are spending more on their websites, digital commerce, and digital advertising than ever before. There has been a shift from traditional advertising to digital advertising. According to a [report conducted by PwC Advisory Services](#), US digital ad revenue increased 16.9% from 2018 to 2019 driven by Connected TV (CTV) advertising. There is a high need for media and broadcasting solutions and services for advertisement and data management and this is driving the demand for digital & media asset management solutions.



**OVER HALF
THE GLOBE
(ALMOST
4.57 BILLION
PEOPLE)
HAVE ACCESS TO
THE INTERNET**

Why is digital & media asset management important to your organization?

Traditional digital & media asset management (D/MAM) tools are used to dealing with static files such as images. They have limited ability to handle large media assets like audio and video files. Legacy D/MAM tools struggle with the ever-increasing volume of audio and video assets that need to be handled daily.

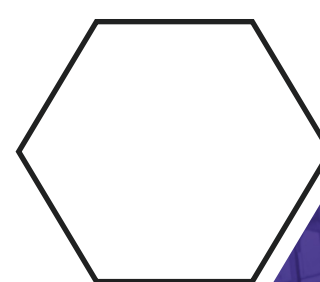
Today, a new breed of tools are available with advanced features – suited to multimedia use cases – using artificial intelligence, machine learning and speech recognition technology to extract advanced metadata information. This metadata is integral to getting the most out of modern video and audio assets.

A 2015 [study](#) by IDC found that 76% of people said that digital asset management (DAM) makes it easier to find assets. This significantly reduces the time spent recreating assets that already exist but cannot be surfaced.

To cope with the increasing volume of assets, organizations need more

“The GrayMeta Curio Platform uses a range of best of breed machine learning services, including Speechmatics, to help media and entertainment companies save time & costs and find new ways to monetise content. For example, we use the Speechmatics service to automatically generate a transcription so that editors can quickly find references to different characters in the video timeline when they’re working in the GrayMeta Adobe Creative Cloud panel.”

MATT EATON, MANAGING
DIRECTOR EMEA,
GRAYMETA



information to identify what makes each asset different. For example, what they contain, keywords, themes, content, contributors and so on. It’s this information that makes assets easier to locate and quickly find out what is contained within an asset.

While there is huge value in automating and augmenting the role of humans by using automated technology, there is still a requirement for humans to be involved in the process of D/MAM. Automated technologies – such as voice technology – enables businesses to manage the volume of digital and media assets cohesively. Giving Asset Managers and Producers access to the right automated technologies lets them focus on curating the best content in the least amount of time, rather than spending hours searching for what they need.

“Automatic speech recognition technology provides significant value to Veritone aiWARE and our portfolio of AI-enabled applications by taking unstructured voice data and turning it into actionable insights. By partnering with innovative cognitive services, Veritone enables more efficiency for our customers by delivering highly accurate speech-to-text, and high-quality, real-time transcription functionality.”

GREG LOOSE, HEAD
OF MEDIA &
ENTERTAINMENT,
VERITONE

The power of voice technology

Voice technology enables organizations to take any media content – whether audio or video – and convert it into text. Converting voice within media content into a written format enables businesses to utilize raw material to then create metadata tags, curate social media content, add captions, search for keywords within an asset and so on. This can be achieved consistently and at scale.

While users can extract information such as the time the file was created, by who, using what software, format, duration and file type, this doesn't tell you anything about the actual contents of the file. Using voice technology, users can expose a deeper context around the content of the file. An accompanying text file to a media file is small in terms of data size (especially HD video). Voice technology makes all the voice data within an audio or video file visible to help content Producers, Editors or other video professionals find the elements that they need in a simple search.

To deliver value and revenue from media assets, organizations need to be able to quickly and easily locate the assets. Automatically transcribing audio and video files using voice technology means all assets are searchable via text-based search. The use of voice technology for the automatic transcription of media assets means that large volumes of assets can be better understood and indexed, making them easier to discover, edit and share on digital channels.

"The ASR technology we have integrated into Tedral Evolution MAM and SMARTLIVE auto highlight creation is a fantastic tool which provides a real value add to our customers.

It enriches the content with information about what a commentator says or provides automatic transcripts of interviews. This information – once in the MAM – can be easily searched and cut within seconds into an edit. Before ASR, this workflow would have required manual transcription which added a significant cost and time overhead."

DAVID ABEL, SALES
DIRECTOR, TEDIAL

Voice technology provides significant efficiencies in the production workflow for digital & media asset management in organizations. This guide will focus on some of the benefits to businesses when using voice technology for digital & media asset management. These benefits are outlined to the right.

- 1 Speed and accuracy
- 2 Time and cost optimization
- 3 Content unification
- 4 Enriched metadata
- 5 Effective archiving

Getting started

Audio and video assets vs image assets

Digital & media asset management covers a range of media types, from images to video and audio assets. Image management is an important part of the asset management process, however, video is fast becoming the most significant media asset to organizations. Video assets have a significant part to play when it comes to social media, lead generation and revenue and this is only growing in the current economic landscape, with most businesses around the world shifting to a digital-first strategy.

67% of marketers find Facebook to be the most important social media channel for their business. While this covers mainly business-to-consumer (B2C) there is an element of business-to-business (B2B). The way that organizations use social media platforms like Facebook to engage with prospects has changed. While images remain important, video is the medium of choice because it delivers the best quality engagement and value. For this reason, organizations with a large number of video assets are looking to their existing assets to create real value before creating new content.

67% OF MARKETERS FIND FACEBOOK TO BE THE MOST IMPORTANT SOCIAL MEDIA CHANNEL FOR THEIR BUSINESS

"80% of users recall a video ad they viewed in the past 30 days."
FORBES FROM HUBSPOT

At a glance

A smart guide on the benefits of using voice technology for digital & media asset management within your business.

Speed and accuracy

Content creators need to be able to download, search, edit and curate their content almost instantly. Voice technology improves this process significantly at scale.



1

Content unification

Organizations can utilize voice technology to more easily search for and locate all of their media assets in one central depository whether it be digital, advertising or live broadcast video files.



3

Effective archiving

Media asset archives are a goldmine for legacy content that is inaccessible and unusable without a transcript to provide enriched metadata. By transcribing voice within a media file into text, organizations can bring legacy media files back to life.



5

Time and cost optimization

By utilizing voice technology to transcribe media assets into a text-based format, editors and producers can search for and store their files faster than ever before.



2

Enriched metadata

Metadata is key to better customer experiences on OTT platforms but unless a video or audio asset has an accompanying transcript with it, the metadata is limiting. Voice technology can unlock the future of enriched consumer experiences with video content.



4

The benefits of using voice technology for digital & media asset management

1 Speed and accuracy

For media asset management (MAM), speed and accuracy go hand in hand. Content creators in markets such as sports broadcast need to be able to download, search, edit and curate their content for analysis, promotion or publicizing on social media or other online formats immediately after an event has happened. As the volume of broadcast content increases, so

too does the demand for curated content. Accuracy of speech-to-text transcription is essential for quick search and captioning of online video content – often involving highlight edits, clips or full programs online.

Using any-context speech recognition technology to automate the process of turning audio or video into a text-based

format significantly speeds up the production workflow for the content curation team. Automating a laborious task such as sifting through hours' worth of broadcast content for online highlights means editors and producers can focus on the creativity and quality of their content rather than manually searching for the specific clips they need.

In the media and broadcast market 100% accuracy is required for most use cases, otherwise, the asset remains difficult to search, or the broadcast presents embarrassing blunders. For this reason, human transcription is necessary to obtain this level of accuracy. Human transcription is time-consuming and

expensive, often taking up to 4-hours to transcribe 1-hour of audio. With growing volumes of content created, the market is looking to speed up the transcription process and offset the associated costs.

Organizations are using voice technology to do the heavy lifting when it comes to transcribing voice within video and audio files. Where humans take up to 4x the length of the audio file to transcribe it, speech-to-text providers deliver 0.5x. Once transcribed, humans can be used for the value-adding editing process to ensure 100% accuracy is delivered.

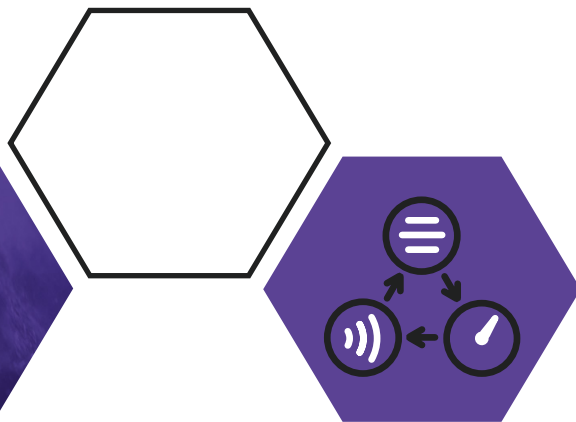
2 Time and cost optimization

Organizations with content are looking to reduce the time and cost related to the supply chain of their media assets. They are focused on outcomes and the ability to extract information from their content to help speed up and streamline their production workflows.

Voice technology enables organizations to automate the laborious part of the media asset management process. Automation allows Editors and Producers to focus their time on actually editing and producing the content rather than sifting through assets searching for the clips they need. Transcribing voice from audio and video files into text helps to reduce the time spent searching for content from live broadcasts or archived content. Content

creators can simply search for keywords, themes or other elements within the file, providing significant cost optimizations.

Archiving media files is notoriously expensive due to associated storage costs. Once archived, files are often left unusable. Transcribing media files into text requires a fraction of the storage capacity and enables creators to locate their content more efficiently. Not only does this help with the efficiency of the production workflow, but it also improves the accessibility of those workflows. With one depository and the right tools available, individuals working from multiple locations can access and use all available content.



3 Content unification

Digital & media asset management is often used by sports organizations. This works as a good example for the complexities of how asset management is vital for any organization with large volumes of media assets.

These organizations are often dealing with a whole host of media assets daily including digital, advertising, marketing, PR, live broadcast and raw content capture. For example, in motorsport racing, it could be many different organizations that provide these services to the rights holder. This could mean that the content and footage remain in different places across the service providers and is hard to share and amalgamate into one central depository.

The goal of a media asset management platform is to aggregate each media asset – regardless of its original format – and make it available to anyone that might need access to it. This list could include many organizations and even the press, journalists, sponsors and broadcasters. Rights owners will often sell the content to broadcasters ahead of time, so speed and unification of media assets are essential to enabling better monetization of assets. But most importantly, content unification of media assets is critical so assets can be found quickly and easily and can be used for an intended purpose.

Voice technology automatically turns voice data within any video or audio file into text and makes searching for any file and specific elements within them a quick and easy task. Ensuring all assets can be searched for in one depository using text-based search means no asset will be missed in the search process.



4 Enriched metadata

Historically, asset management focused specifically on images and other static media. While these elements remain popular, video has taken the mantle regarding its creation, management and value to organizations. While there is certainly room for both video and images, often there is little competition due to the very different use cases and audiences each medium serve.

Asset Managers have now been forced to look at the tools and processes that allow them to manage all assets effectively. Many organizations have neglected video and audio file asset management due to legacy tools and processes in place. This has led to huge archives of files that lack metadata and any sort of value.

Now, with online video content growing rapidly, there has to be a focus on video assets as well as static assets. Organizations are looking to automation technologies as well as artificial intelligence (AI) and machine learning (ML) capabilities to enrich their production workflows to also include video and to optimize the value of these assets to both themselves and their customers.

Metadata is responsible for powering recommendation engines and helping end users discover content that is relevant for them. Engaged audiences on OTT platforms are driven through metadata on the video files. The ability for companies to personalize the experience for their customers is powerful. Metadata tags enable OTT services to do a better job of recommending content and syncing up users' preferences to the metadata tags as they become richer in data.

This process is currently being done by humans through manual metadata entry, however, with the introduction of voice technology into the workflow, organizations can develop content archives and files enriched with more useful metadata. Unless a video or audio asset has an accompanying text-based transcript of its contents, it is almost uneconomical for an organization to gain any metadata value from the original file.



5 Effective archiving

Organizations want the ability to better understand their legacy media asset archives and to treat them much like their contemporary archives which are rich with metadata and value. This level of value can only be achieved through the use of voice technology to convert the contents of assets into text and for a transcript to be stored with all audio and video files.

Some companies will have decades' worth of video and audio media assets that are inaccessible and unusable due to the limited metadata available to search and locate any one asset. Without enriched metadata, humans would be required to listen to or watch all video files based on just file name or date. This process is highly laborious and time-consuming and often leads to content being recreated.

By introducing voice technology to transcribe legacy archives, the benefits can be tenfold. Old, legacy media assets can be monetized, especially for media and broadcast organizations. The files themselves can be brought in line with contemporary files and can be used quickly and easily. By exposing the value from within the files, organizations can bring old content to life and repurpose ideas, but this is only possible by improving the usefulness of files using a transcript and enriched metadata.



Conclusion

Being able to quickly search, locate, edit, curate and monetize video and audio media assets at scale is essential with the growing volume of content now available. With demand always on the rise, Asset Managers, Editors and Producers need to be able to curate content quickly to keep up with the ever-changing digital world.

Between newly created content and legacy archives, organizations need to automate their processes to derive value from them. Without turning the contents of audio and video assets into text, companies are limited to what they can achieve within budgets and timeframes. Voice technology enables businesses to automatically transform voice within audio and video files into text quickly at scale. The result? Searchable media assets enriched in metadata.

With enriched metadata, organizations can start to use insights to drive better customer experiences on OTT platforms. Companies can also drive better consumer engagement with content on social media and other digital channels. The power of digital communication is only getting stronger and customer expectations are following a similar trajectory. Businesses need to be investing in tools and processes to ensure they stay on

top of the volume of assets that need to be curated each day to appease an expectant customer.

The value of metadata rich, usable archives that can be searched and located easily should not be underestimated. This goes hand in hand with the need to turn content around almost instantly from live broadcast to post-event analysis and highlights on a social media feed.

**IN AN INCREASINGLY
CONNECTED WORLD,
CONSUMERS NEED TO STAY
UNITED THROUGH AN 'ALWAYS-
ON' MEDIA CULTURE. NOW, MORE
THAN EVER – IT'S TIME TO AMPLIFY
YOUR DIGITAL & MEDIA ASSET
MANAGEMENT USING VOICE
TECHNOLOGY.**



