

The
NOW
WHAT
Issue

*A new remote-work reality will
mean long-term changes for how
M&E gets business done.
Welcome to Hollywood's
new normal. P. 15*

M**AND****E**
JOURNAL
Media & Entertainment
Strategies. Solutions.

Workflows A New Work Mindset P. 34

Smart Content The Power of AI P. 86

Security Locking Down Your Remote Workforce P. 100

Working From Home What Now? What Next? P. 128



Accomplishing Rapid Closed-Captioning Across All Platforms, Using AI

Demand is increasing for captioning across a much larger volume of media, with many different specifications for caption style, format and language

By Russell Wise, SVP, Digital Nirvana

Abstract: Artificial intelligence (AI) and machine learning (ML) processing in the cloud today offers impressive accuracy, scalability and trainability across a variety of applications useful to production houses and content creators. To benefit fully from these services, however, users must be able to integrate AI- and ML-driven microservices into the on-premises workflows that already support day-to-day operations.

New OTT platforms are entering the streaming marketplace at a dizzying speed, vying to address consumers' continually increasing demand for internet-delivered media. The global movie and TV production communities have taken note and are scrambling to supply content, both new and repurposed.

As these content owners rush to prepare programs and movies for streaming platforms, they must address the unique content submission requirements for each platform. These requirements go well beyond encoding profiles to include various types of metadata and how they are handled.

As a critical subset of this metadata, closed captions are subject to specific rules dictated both by the streaming platform and by its target audiences and geographic regions. Accurate captioning adds value to content and increases its utility for hearing impaired audiences, as well as a broader group of viewers who rely on the service to make content more accessible. Content providers thus face increasing demand for captioning across a much larger volume of media, as well as the challenge of addressing many different specifications for caption style, format and language.

By integrating, automating and orchestrating closed-captioning as part of the overall broadcast workflow, it is possible to take advantage of flexible, intelligent cloud-based STT engines while working within an intuitive user interface that streamlines delivery of compliant, conforming high-quality captions.

While cloud-based speech-to-text (STT) services for closed-caption generation and subtitling are a viable option, managing these processes in-house can be time-consuming. By integrating, automating and orchestrating closed-captioning as part of the overall broadcast workflow, it is possible to take advantage of flexible, intelligent cloud-based STT engines while working within an intuitive user interface that streamlines delivery of compliant, conforming high-quality captions. This model allows for efficiency in creating closed captions, verifying that they conform with style guide from streaming media platforms, and creating translations following the same guidelines. It not only allows content providers to meet accelerated delivery timelines, but also affords closed-captioning houses the functionality they need to offer rapid turnaround times to their clients.

Logical workflow

For either type of user, the workflow is straightforward. Media files tagged for captioning first are transferred via API or portal from a MAM, PAM or storage system (cloud or local) into a captioning watch folder. These files are automatically transcoded into the proxy format used by the host cloud platform and STT processing engine. Because STT processing and other capabilities can be accessed as mi-

croservices in the cloud, uploaded proxies of media files can easily be run through a translation engine as well to generate transcripts in one or more additional languages.

Following processing in the cloud, both a transcript and time code data are delivered back to the on-premise system, where they are available via the internal MAM and through the captioning system UI. Today's highly trained STT engines provide exceptionally accurate transcripts, and custom dictionaries and program-specific information (names, slang, jargon, etc.) can be used to yield even better results.

Converted into captions, the text is shown alongside time-synched video in the review interface, where the operator can quickly go from marker to marker, or line by line, to correct or enhance the copy with visual cues and tags. Automated error detection aids in quick identification and resolution of any errors. During this process, the operator also can review any additional metadata — existing or added following cloud-based processing — associated with the content. This capability offers content providers with an easy way to incorporate content classification into the captioning workflow if they wish and to add further value for downstream processes.

Following review and correction, presets for each target streaming platform are applied to ensure that the new captions

conform with the appropriate captioning styles, which cover number of characters, number of lines, caption placement and other elements. Encoding profiles are likewise applied to generate outputs in the formats required for distribution.

By automating critical steps throughout the captioning workflow and by leveraging AI and machine learning in the cloud, this approach empowers content providers to prepare media for streaming platforms quickly and with confidence. Eliminating the need for human intervention wherever possible and supporting fast decision-making when the human touch is needed, this model allows a single operator to review, enhance, and approve captions with unprecedented speed.

Orchestration of the whole process allows the content provider to centralize management of the end-to-end workflow by assigning jobs and roles, overseeing caption edits, generating reports, and monitoring the status of each captioning project.

The rise of streaming media has brought new relevance and value to the media archives maintained by broadcasters, movie studios and other content producers/owners. Consumers are demonstrating a desire for content on new platforms, and the key to success lies in being able to address this demand — and the technical requirements of compliant content delivery — as quickly and cost-effectively as possible. Offering essential flexibility and efficiency, highly automated captioning workflows driven by cloud-based microservices are a uniquely compelling solution for today's content providers. ■



Russell Wise brings more than 25 years of industry experience to his role, where he focuses on building sales of the company's compliance-logging, intelligent closed-caption generation and metadata enhancement solutions portfolio. russellwise@digital-nirvana.com @Digital_Nirvana