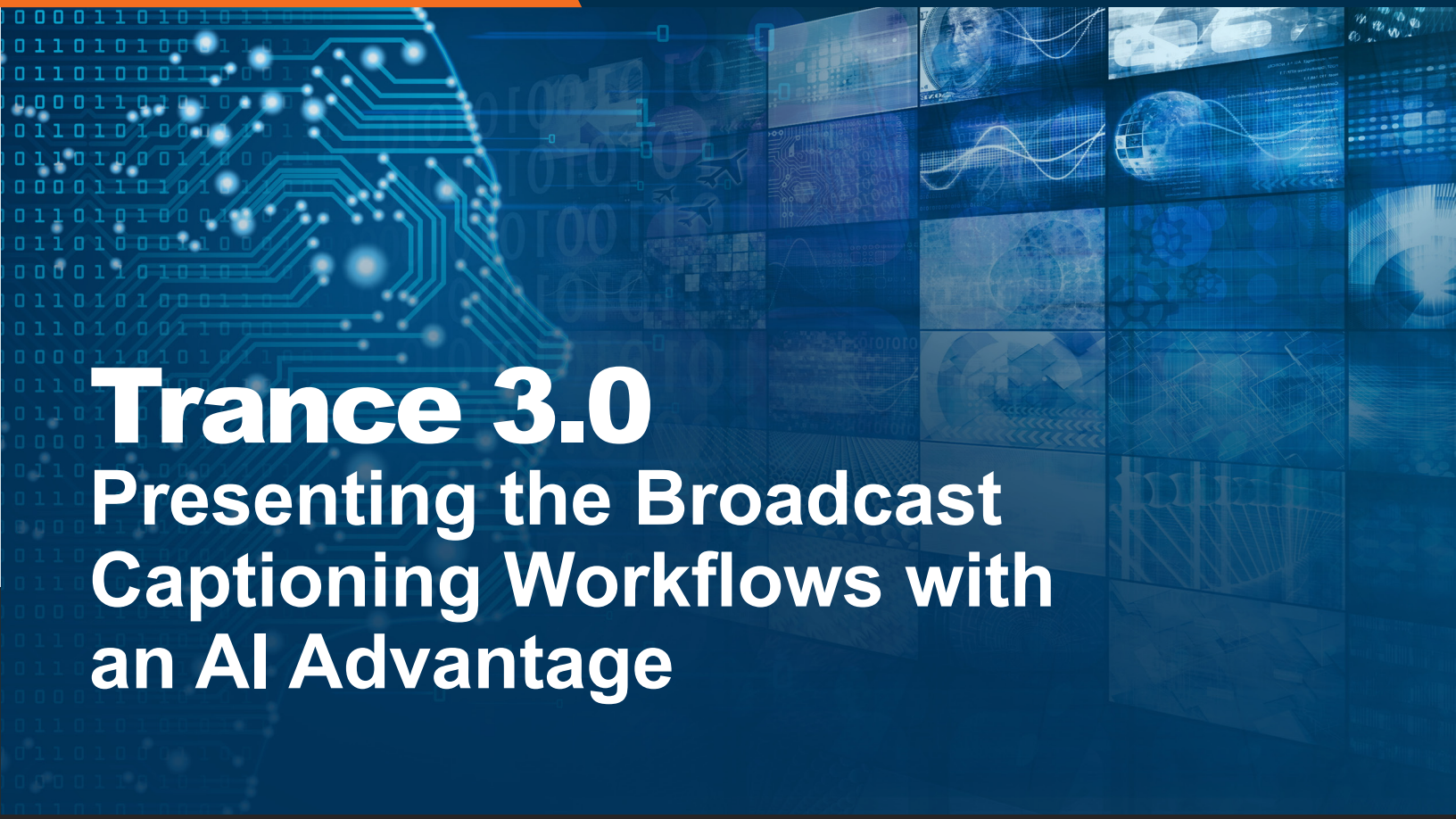
digital nirvana

WHITE PAPER



Trance 3.0

Presenting the Broadcast Captioning Workflows with an AI Advantage

www.digital-nirvana.com

Introduction

Cloud-based services driven by artificial intelligence (AI) and machine learning (ML) offer broadcasters and other media companies the promise of accuracy, scalability, and trainability across a variety of applications. These microservices also yield valuable cost savings and enable a greater degree of flexibility in adapting to changing business and workflow requirements. Supporting agile media processing on a much greater scale than ever before possible, they open up new opportunities for content repurposing, customization, and distribution to different platforms worldwide.

While the benefits of cloud-based media processing are fairly well accepted across the industry, broadcasters still have relatively few options if they want to build microservices seamlessly into their existing broadcast workflows. This is particularly true for closed-captioning and metadata enhancement workflows, both of which are critical to effective monetization of content for broadcast or OTT services.

Many broadcasters do take advantage of cloud-based speech-to-text (STT) services for closed-caption generation and subtitling, and advances in STT technology are introducing unprecedented levels of accuracy. Current STT engines have been trained and refined to the point that they can handle countless words and a wide array of accents, generating remarkably accurate captions.

But even with access to this technology, most broadcasters find the end-to-end process of leveraging cloud-based services too time-consuming and resource-intensive to perform in-house. Staff must have enough technical know-how to access content, create and send a proxy version to the cloud for STT processing, and then retrieve and integrate caption data into the media asset management (MAM) system. Manual correction and formatting of captions demand a further investment of time adding up minutes as captioning for each version of content is adapted and output is delivered in the format appropriate for the target distribution platform.

Broadcasters can choose to outsource closed-caption generation and thereby eliminate some of these tedious time and resource-consuming processes. They stand to gain far greater efficiency if they can integrate, automate, and orchestrate closed-captioning as part of their own overall broadcast workflow. For this approach to work, broadcasters need not only the best-in-class AI-driven services but also the functionality essential to manage those services as an integral part of day-to-day operations. Given the pace at which the industry is evolving, broadcasters also need to know they are working with a solution that can adapt over time and allow them to leverage the latest and greatest in cloud-based processing technology.

Solution:

Trance is Digital Nirvana's enterprise-level, a cloud-based web application for closed-captioning and metadata enhancement solution, designed to empower fast and accurate caption generation. Leveraging two decades of experience in providing services to the media industry, Digital Nirvana created the platform as an internal tool for providing customers with exceptional closed-captioning, metadata generation, and transcription services.

Refined through relentless efforts to improve productivity and contain costs, Trance has exponentially increased the productivity of Digital Nirvana's in-house teams. Now broadcasters around the world are taking advantage of the platform to achieve similarly significant time and cost savings. Trance's cloud-based architecture, the select combination of processing engines, and intuitive interface are bringing significant new efficiencies to the closed-captioning and metadata-generation workflows of broadcasters and media groups.

The Platform

The Trance platform delivers AI-powered cloud-based microservices and an enterprise-level workflow to support highly efficient closed-captioning, metadata enhancement, quality assurance and conformance, and content classification.

The solution deploys very quickly, interfacing smoothly with existing systems and workflows via REST API. Since it can connect directly to an installed Avid system, Trance can automatically extract assets from Avid Media Central (as well as other cloud and local storage systems) for processing and enhancement. As a cloud-based system, Trance is maintained and updated remotely by Digital Nirvana to ensure the utmost in security and functionality. The system's browser-based interface can be accessed remotely by authorized staff working from any internet-connected site, allowing for a collaborative workflow without geographic constraints.

The Applications

Trance ensures that key cloud-based processes can be leveraged with a minimum of technical expertise and human intervention. It eliminates the need for manual movement and transcoding of content throughout the captioning and metadata generation workflow, allowing users to focus on delivering other significant outputs.

Metadata Generation and Enhancement

Trance empowers broadcasters to create rich metadata using STT and video intelligence for video assets. Content can be processed with engines dedicated to face recognition, emotion recognition, and object ID, along with STT. The wealth of data generated by these engines can be paired with timecode data to enable accurate and detailed content searches.

Trance automates the movement of content to and from the cloud for processing and ensures that newly generated metadata is ingested into Avid or other MAMs. Showing video and time synced metadata side by side, the Trance interface makes it easy to examine and review metadata associated with the video.

Effectively working with content that has been indexed with markers and text, the broadcaster can quickly search through hundreds of hours of media to find just the clip needed to create a more engaging program — or to personalize different versions for different audiences.

Closed-Captions/Subtitle Generation

Trance automates the captioning workflow by automatically transcoding files in a watch folder into the proxy format best suited to the host cloud platform and STT processing engine. The process starts with the extraction of media files from MAM/PAM. Content is transcoded and sent to a cloud-based Speech-to-Text Engine. The results are displayed in a single easy to use interface for content editing by in-house captioners. The user interface shows time-synced video and captions side by side in a window along with tools for editing text and adding visual cues, music tags, and speaker tags. When dictionaries, scripts, rosters, and other text resources are ingested into Trance, they can help to boost the accuracy of a transcript and, ultimately, the closed-captions applied to video.

Users are immediately notified of possible errors by Trance's built-in automatic error detection function. The system highlights each instance and allows the user to move quickly from marker to marker or line by line to correct or enhance the copy as needed. This model of automated STT processing can facilitate a closed-caption generation workflow that is 50% more efficient than working with conventional desktop applications.

Upon completion of editing of the STT transcript and addition of all cues, the users will be able to generate captions automatically based on presets of a style guide. These presets include character limit, line length, text frame gap, etc. Based on the presets the application will generate time synced caption and display it in a Pro Captioning Editing window for review.

Since Trance leverages cloud-based microservices, it also can provide a translated version of the transcript. In this case, multiple transcripts are automatically generated, ingested into the system, and then presented for review, comparison, and correction in a dual-pane windowed display where users can see the source language and the automatic translation side by side. Spell-check and error alerts are based on the dictionaries, guidelines, and style rules appropriate for each language.

Closed-Caption Conformance and Quality Assurance

While the accuracy of STT processing is critical to the quality of closed-captioning, the formatting of captions is equally important. Every distribution platform — broadcast and OTT — has a style guide that dictates how closed-captions should be presented, and the content will be rejected if captioning doesn't match up to this style.

Trance, therefore, allows users to configure presets for each platform — Netflix, Hulu, and a host of others — so that caption formatting can be both automated and accelerated. Specifying caption characteristics such as number of characters, number of lines, and caption placement, each preset ensures that closed-captioning conforms with the rules of the target platform.

Content Classification

Trance provides straightforward content classification tools that make it easy to access, view and apply consistent labels to content. Metadata generated by various engines —STT, face and object recognition, etc. — can support fine-tuned content classification, facilitating the automatic generation of labels.

Drop-down menus provide the user with a preconfigured and uniform selection of labels, ensuring that content is classified consistently and according to a scheme that optimizes subsequent search processes. Authorized users can not only ingest and modify existing labels but also add or edit new classification labels. Trance supports the export of classification labels in a variety of formats.

Fine-tuned content classification allows intelligent ad recommendation engines to place content more accurately. As a result, broadcasters, cable operators, and OTT platforms can offer advertisers enhanced targeting while providing viewers with ad content that aligns more closely with their interests or needs.

Orchestration and Project Management

Trance orchestrates the metadata generation and closed-captioning processes from end to end and provides a project management layer that streamlines critical tasks. The platform centralizes workflows such as job allocation, caption edits, role assignment, report generation, and status monitoring. Performing a variety of back-end processes such as file transcoding and transport, the integration of data into the MAM, and more, it also minimizes the need for manual oversight and intervention.

The project manager simply configures roles and priorities for different users and then sets up individual projects by identifying necessary tasks, outputs, and deadlines. Trance automatically handles the movement and processing of both content and metadata; tracks the productivity, workload, and availability of different staff members; identifies the most appropriate person to task with a particular job, and delivers notifications and alerts as needed to drive each project through to completion. In addition to automating scheduling and routing, Trance generates project- and user-based reports that reflect productivity and efficiency.

Conclusion

With a solution such as Trance, broadcasters gain the benefits of AI-driven processing within the familiar media workflows that drive their business. Implementing cloud-based metadata generation and closed-captioning as part of their existing operations, broadcasters can radically reduce the time and cost of delivering accurate, compliant content for publishing worldwide. They can enrich and classify content, enabling more effective repurposing of media libraries and facilitating more intelligent targeting of advertising spots.

Orchestration, automation, and project management tools boost agility and overall efficiency as broadcasters and other media companies deliver content across new platforms and markets. The cloud-based architecture of this model ensures they can take advantage of future technological advancements to address changing market and distributor requirements adeptly.