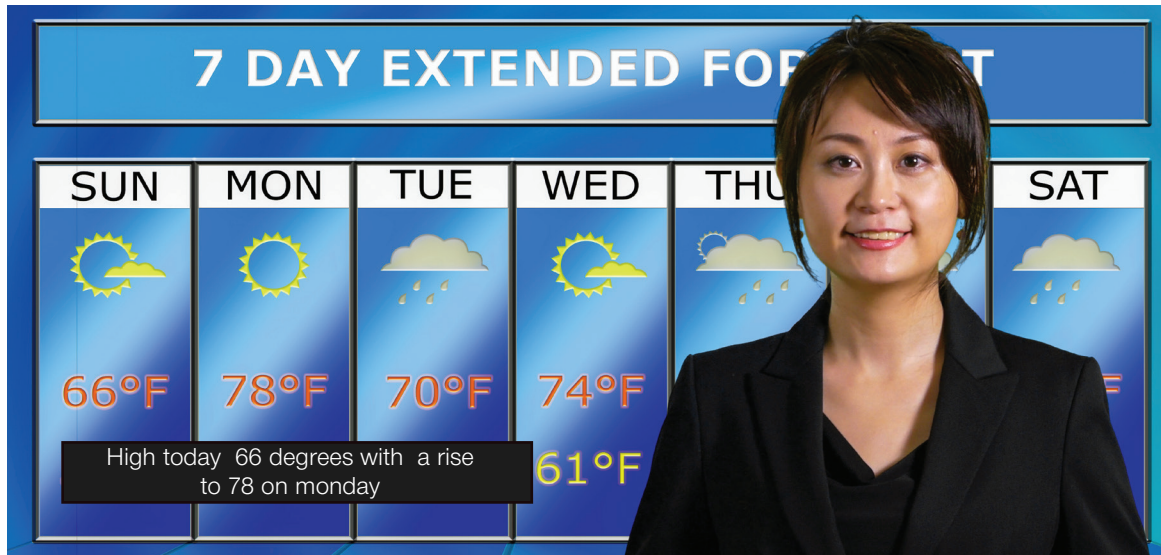


Telestream (Mostly) Automated Captioning

Solution Brief



Next Generation Captioning for Live to VOD Productions – most steps are automated

Live productions such as sports, cultural or political events, are increasingly being produced in a remote or distributed manner with reduced staffing at the event venue and most staff being located at base or even at home.

The cost savings have been shown to be substantial, not to mention the reductions in carbon emissions and improved productivity of equipment and staff. Several recent major sports events and conferences have been successfully produced in this way with no discernible loss of quality.

It is not controversial to say that remote production is now the first choice for most live events.

However, to maximize the audience and revenue from any live event, production companies also want to create Video on Demand versions. These could be highlights packages, news inserts, even full event replays. Whatever the use case, viewers expect these to be available very shortly, in fact almost immediately, after the live event and in the same quality, and ready to view on any kind of device.

While the technology for capturing, editing and repackaging content has been shown to work effectively over the last few major events cycles, one area of the workflow has remained challenging for content producers.

For both regulatory and accessibility reasons, closed captions (or subtitles) must be added to the VoD content. To be acceptable closed captions must be an accurate reflection of the spoken dialogue, and in some regions, these should also include descriptions of the sounds (e.g. “crowd cheering”, “sound of doorbell”) while the economic pressures demand that they are produced very quickly.

Until now, it has been very difficult to meet these two somewhat conflicting demands.



Captioning Technology

For live-to-air events, there are two approaches to captioning:

1. Auto-transcription using Artificial Intelligence
2. Real-time stenography

Stenography is a highly skilled task which can achieve great results but is an expensive process.

AI based transcription is less expensive, but is still prone to errors, especially when the speaker has strong regional accents and dialects or uses slang phrases. Even the best AI transcription will claim a success rate of 95%, which still means that 1 word in 20 could be incorrect. We have all seen some of these transcription errors. Sometimes they are amusing but occasionally they can be very damaging to the media company brand image.

Fortunately, the regulatory authorities allow a little more leeway with live events, knowing that some errors are inevitable.

For Video on Demand, the usual method has been to submit the final recording to captioning staff, who will retype and adjust the captions files for excellent results, even creating translated versions for other languages, and adding the descriptive elements.

While this produces excellent results, it is an expensive process and can add substantial delays to the VOD production process.

The New Semi-Auto Captioning Workflow

Using a combination of existing products, AI transcription and skilled captions editors, Telestream has created a completely new semiautomated captioning workflow, which fits well with cloud-based remote productions and produces high quality captions with the minimum of production delay.

This new solution was tested at a major sporting event, and it proved itself so well that the broadcaster was able to produce well over 1000 hours of accurate captions files over a short period of time, using just a handful of skilled (remote) operators.

Event Live Capture

Live feeds from the event are first captured, using Telestream Lightspeed Live Capture and delivered as “growing files” into a watch folder in a Telestream Vantage system.

The capture equipment is typically located at the venue while the captioning technicians can be located almost anywhere. The capture system is also used to deliver high res and proxy files to the chosen editing platform, with content arriving on the timeline less than 10 seconds “after live”, even when the editing platform is in another continent.

AI-based Live Transcription

The Vantage “Cloud Speech” option extracts audio from the media file, as soon as the recording ends and immediately delivers this to the Telestream cloud-native “Timed Text Speech” service.

Timed Text Speech automatically transcribes the audio to a file with timecode marked text. Timed Text Speech uses the best available AI transcription engines which can work with over 100 different major languages and variants. This process takes just a few minutes even for a long sports event recording.

At this point, the system hands the transcription over to skilled operators for QC or editing. Some (mostly) automated captioning workflows like this have a timer built in where, if a captioning editor does not get to a file in a pre-determined amount of time, it simply moves on to the next step without changes. While this can be problematic with respect to errors in the captions, it keeps the train moving along.

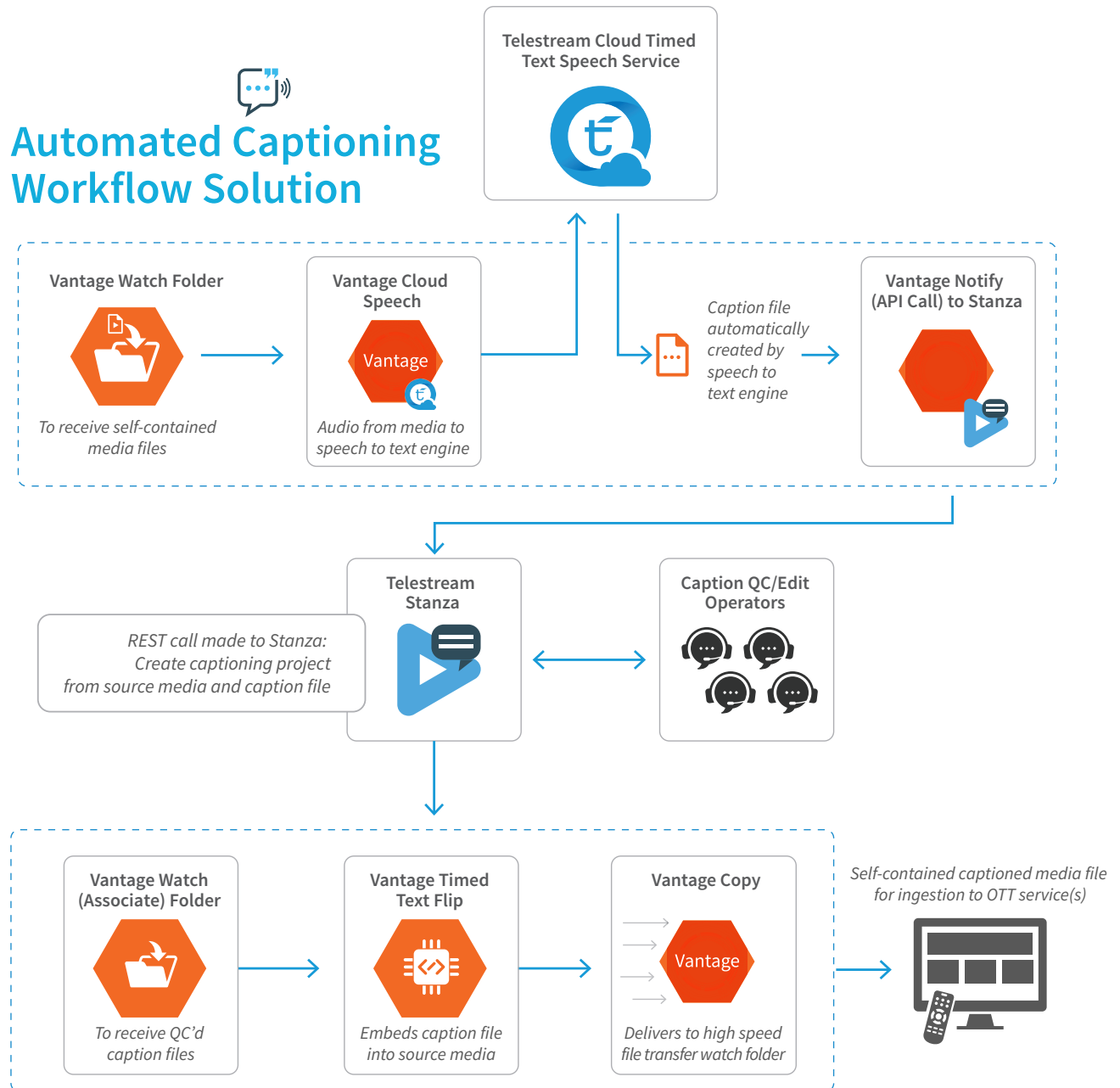
Captions Editing

Once the captions file has been created, the Vantage system sends an API call to Telestream Stanza.

Stanza is a client-server product, which allows remote captions editors to review and edit captions without the delays incurred in copying files out of the cloud, transferring them across networks, or creating proxy viewing copies. Skilled captions editors can work from anywhere because the editing UI for Stanza works from within a browser window. Despite being remote, they can see full resolution frame-accurate video, with multi-channel audio, and can make changes quickly to the captions, either to correct AI errors or to improve readability, or add descriptive elements.

Once the captions editor is happy with the result, they can drop the finished captions file into another Vantage watch folder where it will be converted by Vantage Timed Text Flip into whichever captions or subtitles formats are required for the final deliverable versions.

Automated Captioning Workflow Solution



Final Output

At this stage the edited captions files can be reunited with the captured video to create polished VoD versions which will fully comply with legal and regulatory demands for whichever region and platform the content will be available.

The Final Score

In these capture and captioning use cases, even though the video and captions editors are able to work thousands of miles away from the event, media companies are able to deliver thousands of hours of finished VoD versions with corrected captions, within minutes using this (mostly) automated workflow, while using fewer captions editors than would have been needed using previous technical solutions.

The Telestream remote, (mostly) automated captioning solution combines the speed and automation of AI and cloud computing, while enabling skilled captions editors working from anywhere to polish the final product.

Learn More

To learn more about automating your captioning workflows with Stanza, please follow this link:

<https://www.telestream.net/captioning/stanza.htm>

