

Article Reprint

Captioning's Future

As platforms continue to expand, dubbing's alternative is an economic means to use and resell assets.

By Ed Humphrey, President of the Americas, Softel

Captioning provides an economical means to allow broadcasters and content providers to reuse and resell their assets.

Generally there are two reasons for captioning content: firstly for accessibility – both for the hearing impaired and for when audio is unavailable (e.g. within public spaces

such as at airports, gymnasiums, etc) – and secondly for language translation to enable access to content by foreign language audiences.

Particularly in the digital television, multi-channel world, an increasing amount of content is reused across regions. Captioning provides an economical means to allow broadcasters and content providers to reuse and resell their assets.

Dubbing may make sense in some regions, however in terms of workflow, cost and multi-language access, it does not provide the optimum solution.

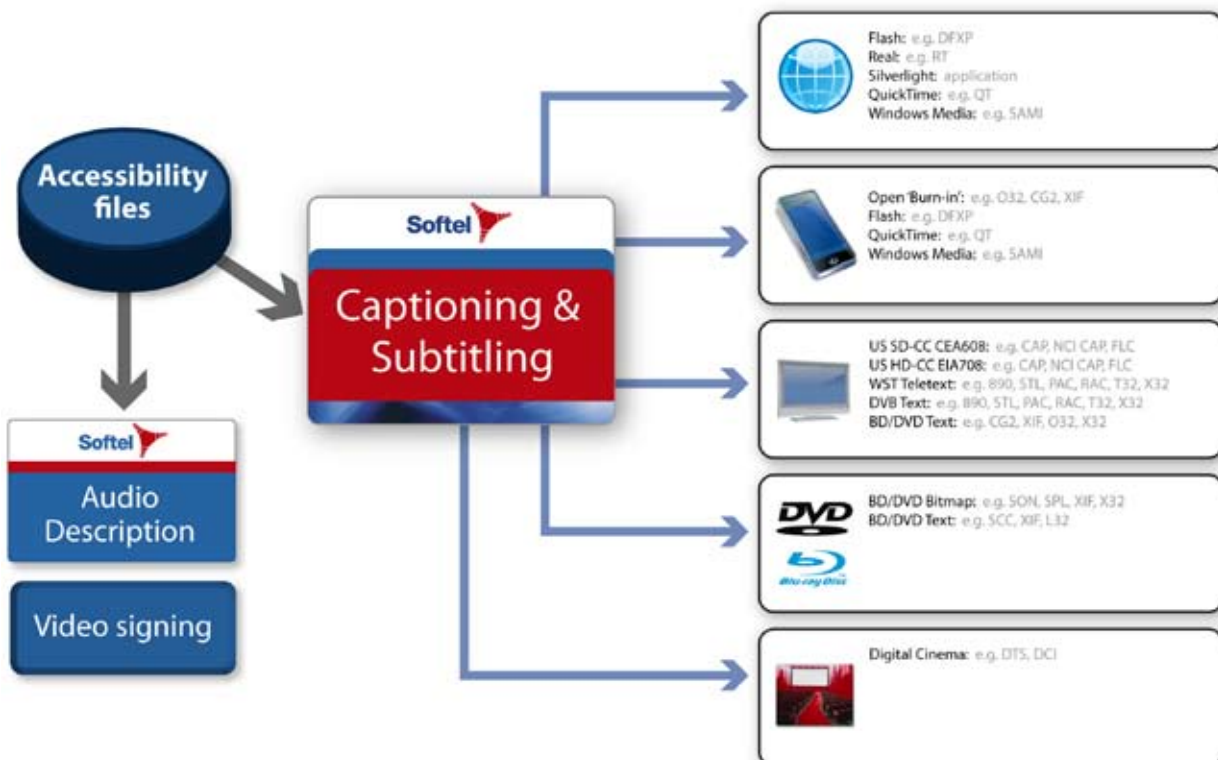


Figure 1 - Deployment of captioning and subtitling to a multitude of platforms

The alternative to dubbing — captioning — provides a number of advantages that make it a highly attractive solution for reaching the greatest number of viewers. The reduced cost makes it more viable for a greater range of content. It is easier and more practical to offer captions utilizing efficient state-of-the-art workflows, especially for content that is live-to-air or has completed production very close to time-of-air. Captions provide the greatest accessibility not just to the main program content, but also to advertisements too, widening the demographic served by the advertiser. Equally, as noted above, captions allow for content to be accessible in public spaces.

Multi-language Captioning

In addition to the necessity to comply with legislation in the target locations, captions provide an inexpensive way to open up content to the wider local and international audiences through language translation.

Dependent on the format of captioning used, it is possible for multiple languages to be delivered to the viewer, and for the viewer to choose their preferred language. In formats such as CEA-708, WST Teletext and DVB, it is also possible to switch on and off captions via the remote control. Figure 1 illustrates some of the platforms and associated formats for captioning and subtitling.

Captions also provide a rich source of content-relevant metadata for the video asset. Enhancing the asset with this additional data aids repurposing, as well as enabling detailed searching and indexing of the content. With integration into centralized media asset management systems, content can also be more easily monetized through clip resale, redistribution and syndication.

Evolving Production and Broadcast Workflows

To aid format and resolution conversions for diverse distribution formats, many broadcasters want to store video assets as a single common ‘mezzanine’ format. This represents the highest quality version. Thereafter, all subsequent broadcast and streaming versions will be derived from it.

To optimize repurposing, the storage of caption data should align with this principle and be stored as a high level generic form of caption data. With this approach there are two key methodologies to consider. One relies on the creation of a ‘master’ caption which has as much information as possible related to the caption, allowing less sophisticated derivatives to be readily produced. In effect this becomes the ‘mezzanine’ format caption and relies on informed choices being made during the creation/preparation phase for presentational aspects such as font, color, positional and alignment information, drop shadow and character edging.

Alternatively, there is the transcode approach, which relies on a lowest common denominator format file (such as an CEA-608 compliant caption or Teletext caption) being created, and effectively ‘up-converted’ to the target format. This will optimize ease of use and speed, but does not take advantage of the sophisticated options available within higher-end standards (such as CEA-708, DVB ETS-300-743 and DVD Bitmap).

Many broadcasters choose to create a hybrid of the two. They implement some of the capabilities while limiting the overall time dedicated to creating the caption by constraining and automating some choices. When considering which approach to take, it is worth noting the variations that different standards offer in terms of levels of control and sophistication.

Captioning for File-based Workflows

The caption creation technologies most commonly in use fully support non-linear video – thereby allowing content providers to digitally send encrypted and/or watermarked video clips. This saves costs associated with tape-based methods.

The best creation systems allow for far greater caption operator productivity through technology enhancements such as shot change detection, semi-automated time-coding functions, advanced reading speed algorithms and so on. Modern caption authoring systems should also be Unicode capable, in order to create and repurpose captions in almost any language. For complete flexibility they also should fully cater for HD, VOD, Web, mobile, digital cinema and for the burgeoning stereoscopic 3D space.

The employment of a wide network of caption operators makes expedited delivery ever more possible. Often these operators are located around the globe in order to maximize time zone benefits. As a result of this highly distributed mode of production, a substantial worldwide network of professional freelance caption operators has emerged, equipped with the latest creation workstations and using high-speed broadband links to receive secured video clips and job instructions.

Finally, creation is sped-up through caption agencies sending “proofs” back to their clients electronically. This is achieved by generating an all-digital “approval”, which has captions overlaid to video so that the client can quickly and easily assess placement, timing, font choice and other factors.

Binding Captions to Content

After the creation phase has been completed, the caption or subtitle data must then be “bound” to the content, enabling presentation to the viewer when they watch the programming.

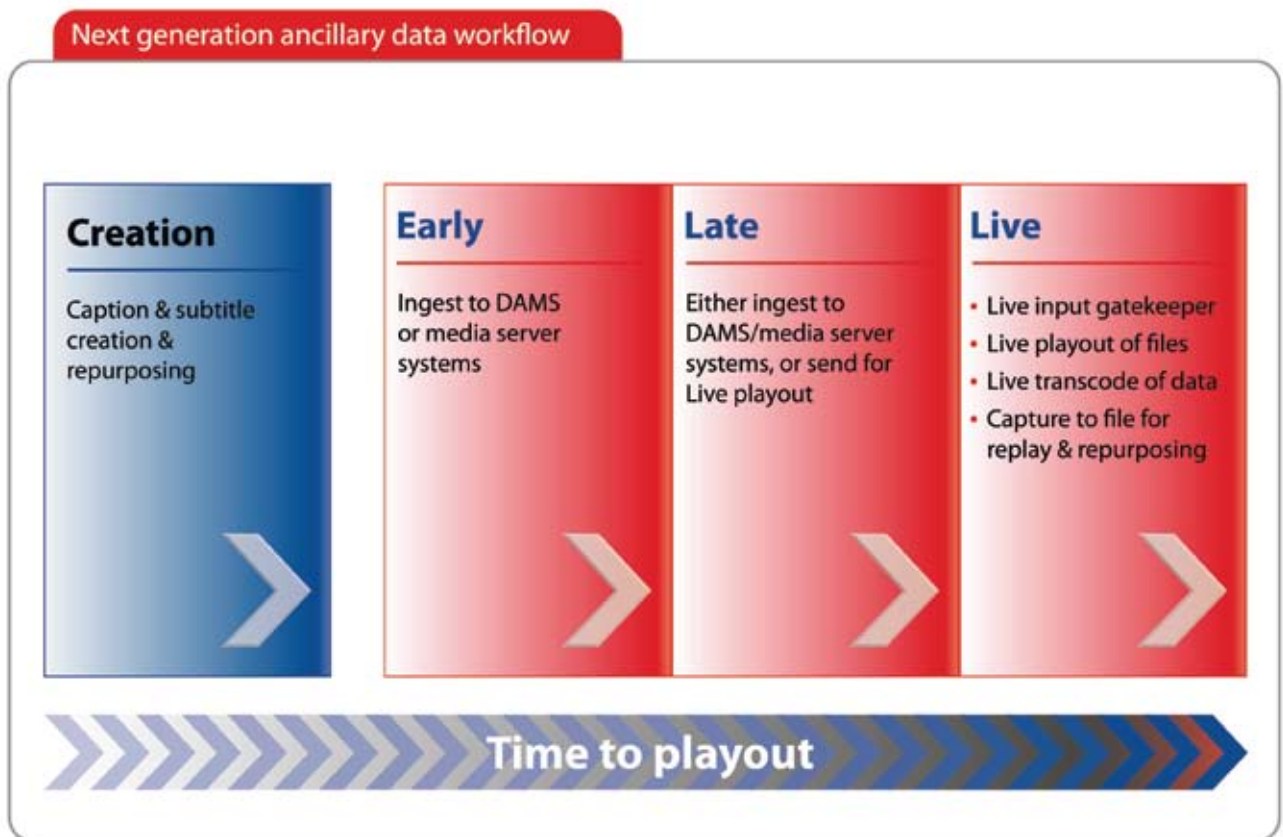


Figure 2 – The caption and ancillary data binding process in next generation workflows

This binding can be considered as occurring in one of three periods of time, as illustrated in figure 2:

- Early binding – The pre-prepared file is linked to the program content well ahead of transmission
- Late binding – Similar to early binding, but occurs closer to air time and only becomes possible due to faster-than-real-time encoding technologies
- Live binding – For either truly live content or for pre-prepared content which only becomes available very close to airing thereby eliminating the possibility of pre-binding captions.

In modern workflows, files are now either sent for time-of-air transmission (a live bind), or are transcoded into a file-based video asset (during early or late binding).

Driving time-of-air transmission will be a system which tightly integrates into the automated workflow of a master-control facility, with the caption playout system approving files in advance of airing, then airing the correct file at the right time automatically, either with or without external time-code.

The time-of-air system can also be used as a 'gate keeper' for real-time captioning, where the system authenticates the

caption operators and their work slot, prior to allowing pass through to air.

Additionally, the time-of-air system can extend the control of the automation system over the live captioning by switching the data source to the in-line caption encoder based on the automation schedule. This removes the dependency on the live caption operator to remember to manually switch control at the right times. This prevents situations such as open connections blocking playout of pre-prepared content during commercial breaks.

Increasingly a hybrid of ingest and time-of-air methods is becoming the workflow of choice, resulting in a system that intelligently arbitrates between ingest to video servers whenever possible and time-of-air playout as appropriate. In this role, the time-of-air playout system is elevated to a central caption management platform.

The time-of-air caption system can also provide interfaces to other ancillary data signals and XDS information such as wide-screen signaling, vChip parental controls, Broadcast Flag information, DRM controls such as CGMS-A data, Digital Program Insertion (DPI) data, etc.

Transcoding Captions for Flexible Asset Management

As well as supporting different output distribution formats, modern captioning solutions support “reversioning” of video assets. This occurs when an asset is manipulated in the time domain or split into different program segments. This “reversioning” process often takes place within a Non-Linear Video Editor (NLE) and can effectively destroy the captioning data, as it becomes disassociated with the video and audio content. Modern transcoding solutions can circumvent this issue by using the Edit Decision List (EDL) from the NLE and other sources of data describing the differences between the original and the derived version of the video, bridging the caption data from the original to target version.

By using the mezzanine format for video and caption data described earlier, the caption data component can be transcoded appropriately at the same time as the video, ensuring the same quality or better of captions as for the broadcast version. The same also applies for situations where the ‘lowest common denominator’ mezzanine route has been followed, which while it may not provide better presentation on playout, still provides the other benefits of easier repurposing and greater efficiency through predictable output.

Seamless Integration in Complex Next Generation Workflows

Captions are an important piece of modern broadcasting, but new thinking and new workflows are essential to keep pace with the roll-out of new platforms, global distribution and evolving standards. Seamlessly integrated automation and efficient workflows are increasingly essential in order to allow for cost effective yet robust delivery to the widest content audience possible.

Any platform for caption playout and management must provide the stability needed to ensure confidence in the data making it to the viewer. Therefore it is vital to seek out technology providers who have the expertise and experience in developing and delivering the specialist systems required, as well as those that have proven their worth in the field. Using this expertise, it becomes possible to implement highly fault tolerant solutions that integrate seamlessly into modern workflows

This article was first published in the September 2011 issue of Broadcast Engineering, USA/Canada edition.

