Captions, Subtitles & Transcription

Definitions

Captioning converts the audio content within a video into text, *then synchronizes the transcribed text to the video*. When the recording is played, that text will be displayed in segments that are timed to align with specific words as they are spoken. Captioning is required to make video content accessible to viewers who are deaf or hard of hearing.

Subtitles show the translation of words spoken in a different language. The words shown on the screen in a foreign film in another language, for example, are considered subtitles.

Transcription is the process of producing a text document from the words spoken in a video. Transcribed text does not have a time value associated with it and can't be used immediately for captions or subtitles - further editing will be required. In terms of accessibility, transcription works well for audio-only media, but falls short when it comes to audio with moving content on a screen, such as voice-over-PowerPoint slides or video.

University Captioning Policy (July 2020)

From 23 September 2020, websites of public sector bodies will need to satisfy new accessibility requirements.

These regulations were brought into law in 2018 as part of the disability-focused Public Sector Bodies (Websites and Mobile Applications) (No. 2) Accessibility Regulations 2018.

Websites that don't meet the new requirements after 23 September 2020 deadline could be found to be breaking the law. This includes intranet services (or internal company websites) as well as websites for external audiences.

Websites will also need to include an Accessibility Statement - sample statements can be found here

FAQ's

There are a few differences between open captioning and closed captioning in videos. Most notably, open captions are always on and in view (burnt in), whereas closed captions can be turned off by the viewer. Open captions are part of the video itself, and closed captions are delivered by the video player or television (via a decoder). And unlike closed captions, open captions may lose quality when a video is encoded and compressed.

There are two important differences between ASR captioning (also referred to as machine-generated captions) and human-generated captions: the quality and the time required to generate captions.

Machine-generated captioning produces captions very quickly. Typically, captions can be created in about one-quarter of the total video length. For example, an hour-long video could be captioned using ASR in approximately 15 minutes.

ASR captions are typically 70-75% accurate depending on the audio quality in the recording. As a result, machine-generated captions are primarily intended to enable inside-video search, and by default, they aren't added to the video as closed captions. Instead, the text is stored in the video platform's database for use with the video search engine.

Of course, ASR also provides a starting point from which people can manually create 100% accurate captions. In video platforms like Panopto, text generated by ASR can be added to the video as closed captions, which people can then edit.

Human-generated captions take substantially longer to produce but provide results that are at least 99% accurate. In some cases, human-generated captions can be turned around in 24 hours, but typically, you can expect a 2-5 day turnaround.

In addition to making video content more accessible to viewers with impaired hearing, captioning can actually improve the effectiveness of video:

- Captions improve comprehension by native and foreign language speakers. An 2006 Ofcom study showed 80% of people who use video captions don't even have a hearing disability.
- Captions help compensate for poor audio quality or background noise within a video.
- Captions make video useful when a person is watching with the sound off or viewing in a noisy environment that obscures the sound.
- Captions provide viewers with one way to search inside of videos. The link gives an example using the University Replay lecture capture.

Machine-generated captions can be generated in a quarter of the time it takes to play an individual video. Depending on the service used to generate captions and the length of material, it can be well within 24 hours.

Human-generated captions are typically generated within two to five days, depending on the requested turnaround time and service options.

Yes, while a live event is taking place captions can be added to the video content but typically this would involve a third party supplier who provides captions generated by machine or human.

The captions can be displayed either superimposed over the video content (typically at the bottom of the screen) or as a separate web browser page so that more than one line of text is visible at a time. The latter is useful for viewers who may need to see captions in a larger font size or need longer to read them if someone is speaking quickly.

It is also possible to provide subtitling (simultaneous translation) in this manner, however **costs are an important factor** in live captioning and translation - it can become expensive.

The latest version of Zoom will generate captions while the event is taking place and individual users can choose whether captions are displayed or not, and what size to make the captions. This facility is free of charge.

The Web Content Accessibility Guidelines, also known as the WCAG standard, is the most detailed and widely adopted guide for creating accessible web content.

WCAG 2.1 AA generally asks that online content meet four principles that improve accessibility for people with disabilities and also adhere to a certain level of compliance. Both are summarized below:

WCAG Design Principles:

- **Perceivable:** All relevant information in your content must be presented in ways the user can perceive.
- **Operable:** Users must be able to operate interface components and navigation successfully.
- **Understandable:** Users must be able to understand both the information in your content and how to operate the user interface.

- **Robust:** Content must be robust enough that it can be interpreted by users, including those using assistive technologies (such as screen readers).

WCAG Compliance Levels for Online Video:

- **Level A:** Captions are provided for all prerecorded audio content in synchronized media, except when the media is a media alternative for text and is clearly labeled as such.
- **Level AA:** In addition to Level A compliance, captions are provided for all live audio content in synchronized media.
- **Level AAA:** In addition to Levels A and AA compliance, sign language interpretation is provided for all prerecorded audio content in synchronized media.

To learn more about WCAG 2.11 guidelines visit W3C.org.

> ⊘ **Automatic Captions using Youtube**
>
> Youtube provides a facility for automatic captioning of videos on it's channels.
>
> Be aware that this only applies to videos which have been recorded and uploaded to the channel - it does not cover live streams. Providing captions (or subtitles) for live streams will involve a third party supplier who will make a charge for their services.

Captions using Panopto

As of August 2020, the University will automatically caption all material recorded by Panopto - the PDLT team have produced a **one page guide** to assist with embedding and editing captions.

In addition to that, there are links below which provide further information:

Downloading captions

Uploading captions

Setting up 3rd party integration - the example uses 3PlayMedia as a provider for caption services.

Adding translated captions

Third party suppliers

AI Media - currently used by the Univeristy's Equality & Diversity team for events and INCLUDE meetings

3PlayMedia - used by Audio Visual & Digital Marketing for one-off captioning requests

Third party suppliers will always make a charge for their services and there may also be costs incurred for cancellation at short notice.

Further Information

There is some **further information here from Digital Marketing & Communications** relating to Youtube captions

Using the central captioning pool to hire and book a student Captioning Assistant