Video Captions and Audio Descriptions

Stephanie Obodda - 2025-12-11 - Comments (0) - Accessibility Concepts

Captions provide a text equivalent of the audio information of a video and are synchronized with the video presentation. Captions are an accessible alternative for individuals who cannot hear the content and includes the dialogue, speaker information, and any sound effects present in the video.

While the terms "captions" and "subtitles" are often used interchangeably, subtitles do not provide the same information and are designed for different purposes. Subtitles provide a text version of the dialogue only and often in different languages. Essentially, subtitles assume an audience can hear the audio but also need the dialogue provided as text.

Tips for captions

- Include all spoken information as well as relevant parts of the soundtrack like background noises, sound effects, speaker identification, and any audio cues that help the viewer understand the video.
- Captions should be in the same language as the primary audio content.
- Use proper punctuation and grammar.

What videos need captions?

- Videos promoting your program or attracting students, participants, and alumni
- Videos showcasing curriculum, research, exhibitions, or collections
- Videos profiling students, faculty, or researchers
- Videos providing instructions for how to apply or register for programs
- Videos listing news stories about your department or program

Zoom and captions

Zoom provides the capability of including captions from a professional captioner or to use auto-generated live captions from automatic speech recognition (ASR) engines. For live events that are not public-facing or in which there is no request for captioning services, follow these steps to enable the live transcription feature in Zoom:

1. As the host, go to Zoom Settings and In Meeting (Advanced).
2. Move to the Closed Captioning option.
3. Check the option "Enable live transcription service to show transcript on the side panel in-meeting."
4. Participants will then be able to choose to display auto-generated live captions during the Zoom meeting or event.

Captioning YouTube Videos

If using YouTube to publish your videos, you can add captions using YouTube's captioning interface. You can start with YouTube's auto-captions to create a draft transcript and then make updates to that transcript to correct for any errors.

Transcripts

A transcript is the same word-for-word content as captions but presented in a separate file. It provides a text alternative of the audio presentation and is not synchronized with the audio timeline. A transcript should contain relevant speaker information to distinguish who is saying what information.

Tips for transcripts

1. While transcripts may be provided for pre-recorded videos, they must be provided for audio-only content.
2. Transcripts should include speaker information or any other informational cues appropriate to understanding the recording.
3. If an audio or video was scripted before production the script can be used as the transcript.

Audio descriptions

Audio descriptions provide a verbal depiction of the key visual elements in a video presentation. For individuals who are blind, low-vision, or unable to view the video directly, audio descriptions communicate the important information relevant to understanding the video content. For example, a video may display a speaker's name and title or specific instructions to follow. If this information is not included as part of the spoken dialogue, then it needs to be communicated as part of a separate audio description.

The Description Key from the Described and Captioned Media Program provides guidance for how to produce audio descriptions, including what to describe and how to describe on-screen information:

- What to describe

- [How to describe visual content](#)

Because audio descriptions exist as a separate track in the video presentation, it can be challenging to include these descriptions once the video is complete. A simpler way is to plan your script to describe any relevant on-screen information.