

Video Accessibility Workflows

Video accessibility can involve several different workflows depending on whether video is being evaluated or created. To ensure that video content is accessible, it is necessary to evaluate both content that the library purchases or subscribes to from outside vendors and content created, preserved, or maintained by the library directly. There are several workflows that can help to ensure accessibility is not overlooked and provide the structure needed for remediating videos that are not currently accessible. These workflows are intended as starting points for this process, though they may need to be refined or modified depending on specific institutional needs.

Evaluating Video Purchases and Subscriptions

When developing workflows around accessibility evaluation for collection development purposes, it is important to include an evaluation of video content in the library's collection. The first step in this process is to request a Voluntary Product Accessibility Template (VPAT) from the vendor, if one has not already been provided. A VPAT is a document that explains how an item, such as a database or piece of software, does or does not satisfy the requirements of a particular accessibility standard. Typical standards that are included are

- Web Content Accessibility Guidelines (WCAG), which is the international standard for web content accessibility;
- Revised Section 508 standards, which govern software and hardware procured by the US federal government and is often used as a standard by other institutions as well; and
- EN 301 549 accessibility requirements, which govern public procurement of ICT products and services across the EU.

VPATs are generally organized by WCAG success criteria, which makes it relatively straightforward to have a process in place for specifically checking video accessibility. The relevant success criteria to focus on for video content are those found in 1.2 Time-Based Media, which covers the requirements for video accessibility to meet Level A, Level AA, and Level AAA conformance levels.¹

Unfortunately, VPATs are not always accurate. A 2015 study of VPATs found an “inaccuracy rate of 19.6%.”² This means that it is worthwhile to make an independent verification of accessibility features part of the evaluation process. While often accessibility evaluations make use of automated testing tools, in the case of video, it isn't possible to fully assess all accessibility features in this way. While these tools can be used for certain elements of the process, as discussed further below, at least at this point, they cannot evaluate the adequacy of captions and audio descriptions. This means that manual verification is necessary to ensure accessibility of video content.

Evaluating Captions

Because captions are integrated into video files, the best way to evaluate captions is by watching the video. For purposes of evaluating a vendor resource, this may mean checking a few videos to confirm that captions are consistently high quality. The following questions can guide this review:

- Are captions present in all videos with sound elements that are integral to understanding the video?
- Are the captions synchronized with the video and its soundtrack?
- Do the captions achieve 99 percent accuracy? If not, estimate how accurate they are to determine adequacy.
- Do the captions indicate who is speaking and whether the speaker is on screen or off screen?

- Are non-dialogue sounds captioned appropriately?
- Can the captions be turned on and off (e.g., closed versus open captions)?
- Are the captions high contrast enough to read over the video?
- Is the font clear, customizable, or both?
- Can the font color of the captions be changed?
- Can the font size of the captions be changed?
- Can the background of the captions be changed?
- Can the captions be moved to another location on the video? If not, does the layout ensure that they do not obscure the video?
- Do the captions censor or otherwise skip important content?
- Overall, are the captions sufficient to allow the user to completely understand the video without the soundtrack?
- Are captions immediately available on new content as it is added to the platform? If not, how quickly is it added, and is there an option to place a request to expedite this process if needed?
- Are the controls for the video (i.e., play, pause, audio levels, toggle for captions) accessible?

Evaluating Transcripts

Because transcripts are not integrated into the video file itself, the process for evaluating their accessibility is a bit different. This will require both an evaluation of the transcript text and an evaluation of the area displaying the transcripts to ensure that both are accessible. Evaluating the area displaying the transcript can largely be tested via automated accessibility testing tools, at least to the same extent as other web content. Transcript text, unfortunately, is less amenable to automated testing tools and needs to be evaluated manually at this point. It also requires checking a few video transcripts to confirm they are consistently high quality. The following questions can help in the evaluation process:

- Does the transcript accurately capture the sound elements in the video?
- Does the transcript include necessary descriptions of key visual elements, represented clearly in a manner so that they are not mistaken for part of the audio track?
- Are there elements that require transcripts, such as sound elements, visual elements, or a combination of the two, that are integral to understanding the video?
- For scrolling or highlighted transcripts, is the motion in synchronization with the video and its soundtrack?
- For interactive transcripts, does searching in or clicking on sections of the transcript move the

- user to the appropriate point in the video?
- Is the transcript in a usable font size and style? Is the font customizable?
- Is the transcript searchable? This feature makes the transcripts more usable for a wider range of users.
- Is the transcript exportable? While this is not absolutely necessary for accessibility, it does make it easier to use the transcript in more ways and for more purposes.
- Is the interface in which the transcript is presented accessible to assistive devices and by keyboard navigation?

Evaluating Audio Descriptions

As with captions and transcripts, it is often necessary to play a video file, or a sample of videos, to evaluate whether audio descriptions are present and whether they are adequate. In some cases, when audio descriptions are listed as a separate audio track or a separate version of the video, it may be clear that the platform offers audio descriptions, but it is still important to manually examine their adequacy. The following questions can guide the evaluation process:

- Are the audio descriptions part of the main audio track or a separate audio track? If the latter, are users able to turn them on or off?
- Are the audio descriptions audible? For audio descriptions that are part of a separate audio track, can the volume for the audio descriptions be adjusted separately from the main audio track?
- Are the audio descriptions at a speed that is comprehensible?
- Do the audio descriptions fit within the natural pauses without overlapping any key elements of the soundtrack?
- Do the audio descriptions adequately convey visual elements in a way that makes the video understandable by those not watching the video?

While it may not be possible to evaluate every single video file included in a platform, this evaluation process can be done with a small sample of videos. If videos are presented in multiple formats, it would be worthwhile to check the different formats as part of this process. As this evaluation is being done, an important piece of the workflow is also documentation. Keeping notes about the results of the review will help in a few ways. First, it makes it possible to offer guidance to users on what is and is not available. Second, it can help when following up to determine whether accessibility has improved or deteriorated. Finally, this evaluation can be made a part of the collection development decision-making process more

easily if there is documentation. It can also be useful when negotiating with a vendor and, when appropriate, the results shared with the vendor as a way of advocating for improved features.

Creating Accessible Video Content

There are many different approaches that libraries can take to incorporating accessibility in videos created in-house, from creating accessibility features internally to outsourcing the work to any one of many different services that caption or describe audio content for a fee. Depending on the nature of the video to be captioned, the time line for creating captions, and the available staff time and skills, different approaches may make more or less sense for a particular institution or project, but these workflows offer options that can be customized for individual institutional needs.

Creating Captions and Transcripts from a Script

One of the easiest ways of creating captions and transcripts is from an existing script. Having an accurate script on hand can streamline the process considerably, but there are still several steps to the workflow:

- Create a script before the video is created, and then record the video.
- Once the video is recorded, correct the script to reflect any deviations from the script during recording.
- Save the script in an appropriate file format. While the exact file formats that will work depend on the platform you are using, SubRip (.srt) and WebVTT (.vtt) are common options that are available across many platforms.
- Upload this file with the video in a platform that supports closed captions, or use video editing software to incorporate open captions into the video.
- In the case of captions or interactive transcripts, check that the file has synchronized properly so that the correct text is displayed at the correct time stamp in the video.

While this process is one of the most efficient ways of adding captions or transcripts to a video, it depends heavily on whether a script has been created and is closely followed in the process of creating the video. This will not be practical in all cases, and, if the script will not be accurate when uploaded, this approach may not necessarily save time in the process.

Editing Automatically Generated Captions and Transcripts

While automatic captions and transcripts are not yet able to reach the accuracy levels needed to provide full access to video content, they can be used as a starting point for creating more accurate captions when a script is not available. This workflow can be used for that process.

- Once the video file is completed, upload it to a service that automatically captions videos. There are many options, including YouTube, Facebook, and Otter.ai. It is important to note that once the video has been uploaded, it can take some time for the automatic captions to be generated. This is generally not an instantaneous process, and the timing can be variable, particularly with free tools, in some cases taking up to several days before captions are generated.
- Assign an individual to review the automatically generated captions. Though this may not seem like a difficult task, it can be time-consuming, especially for those who are new to the process. It tends to be a bit faster when done by the person who created the video or the main speaker in the video, as this streamlines understanding the content in the video. It is also a process where experience can increase speed.
- Review and correct the captions with a focus on the following:
 - *Punctuation*—Often automatic captioning and transcription tools miss important punctuation, and some, such as YouTube, tend not to insert punctuation at all.
 - *Grammar*—Sometimes the speech recognition tools used for this purpose will introduce grammar errors, so it is important to make corrections to ensure that the grammar matches the audio track.
 - *Spelling*—This can be one of the most important aspects of the correction process. Spelling errors will happen most frequently with words that sound very similar to other words, where a proper name is not in the tool's dictionary, when foreign words are used, and where the speech being captioned is accented.
- Add any non-speech sounds that are not included automatically. Generally these are added in square brackets, but some organizations use parentheses. Though square brackets are the best practice, the most important consideration is that these are used consistently within a video and, ideally, across videos at an institution.

- Insert line breaks to ensure that the captions are readable. Generally, a caption should have no more than eight to ten words on a line, though the exact number will depend on word length. Also, it is best to limit the number of lines on the screen at one time so that the captions do not block the video.
- Check and correct timing as necessary. Though automatic captioning tools try to keep the captions synchronous with the video, there may be errors, and it is important to make sure that the captions are synchronized and remain on the screen for the appropriate length of time.
- Once these corrections have been made, save and, if required by the tool being used, publish the corrected captions.
- An optional step in this process is to have another member of the team double-check videos for accuracy. This can be helpful for ensuring accuracy, particularly for those who are new to captioning or transcribing. This process could be applied to all videos, or a few videos could be spot-checked at random.
- A final optional step in this process is to download and archive the finished file so that it can be backed up separately from the platform used to create it (or available for archiving or uploading to other platforms as necessary).

While editing automatically generated captions and transcripts is a significant undertaking, it really cannot be overlooked. Without corrections, these automatically generated texts do not provide the level of accuracy necessary for accessibility. For this reason, it is very important to factor in the staff time required for this process when determining the budget for captions and when deciding which approach to video accessibility the institution will create.

Creating Audio Descriptions

As discussed in chapter 4, the process of creating audio descriptions requires skill and experience. Because they should ideally fit into the natural pauses in the existing audio track and because they require judgments about what content needs to be described, creating audio descriptions is more difficult in some ways than creating captions or a transcript that simply reproduces the exact language spoken in a video. For this reason, it should be expected that the process will take a significant amount of time and will likely include all of the following steps:

- Watch the video in its entirety. Even if the person creating the audio descriptions also created the video, it is worthwhile to watch the entire video with an eye toward which visual elements should

be described and when descriptions will fit. During this first viewing, some notes may be taken, but that may need to wait until a second viewing.

- Once the person creating the audio descriptions has watched the video and taken some initial notes, that same person should be tasked with creating a script of the audio description. This process should be undertaken by the same person who initially started the planning process so they are familiar with the video in its entirety, or, at a minimum, the entire section they are responsible for describing.
- The process of creating this script will likely require viewing segments of the video again and noting the time and length of gaps in the sound track. While the creation of audio descriptions cannot be automated, there are tools that can help with identifying these gaps, such as CADET, discussed in further detail in the previous chapter. The final script should denote the time markers at which the audio descriptions should start and stop.
- Once the script is drafted, it should ideally be reviewed for clarity by a separate party to ensure that it provides meaningful access to all necessary visual content.
- The person tasked with recording the audio descriptions should review the script. The person recording the audio description need not be the same person who created the script, and, in fact, there may be some value in hiring a professional voice-over artist at this point depending on the nature and scope of the process.
- The audio descriptions should be recorded per the timing listed in the script.
- The penultimate step in this process will depend on the platform. If the platform supports a second audio track with audio descriptions, this file can be uploaded at this point. In this scenario, the main audio track would need to be edited only if there was a need to lower background noises or soundtrack elements so they do not obscure the audio descriptions. However, if a separate audio description track is not supported, as is the case in many platforms, the audio description recording will need to be edited into the pauses in the main soundtrack.
- Regardless of the approach taken in the previous step, the final step is confirming that the audio descriptions are properly synchronized with the video.

Because of the divergent skills required to create the script and then record it, this workflow is more likely to involve multiple creators than the others discussed in this chapter. Given the high level of skills involved, the creation of audio description may be an

area where institutions find it more effective to outsource this workflow.

Outsourcing Caption, Transcript, and Audio Description Creation

Because of the time and skill required to create accurate captions, transcripts, and audio descriptions, many organizations opt to outsource the production of these tools rather than creating them in-house. This can save staff time and, in some cases, may even be more budget-friendly, but it is important to note that this still requires a plan and workflow to proceed successfully. While each vendor offers different specific procedures, this workflow demonstrates the basic steps with a focus on where an organization will still need to allocate staff time:

- Once a video file is created, it will be submitted to the selected vendor. There are many ways this submission process can happen, including e-mailing it to the vendor, uploading it to the vendor's website, using an integrated submission feature in another platform, or even integrating it into a project via an API.
- After the vendor receives the video, it will process the video. During this step, the institution should monitor to ensure that the time frame for returning the completed captions, transcript, or audio descriptions is met.
- Completed videos must be manually reviewed for accuracy. Some vendors guarantee specific accuracy levels, but it is still important to ensure that this accuracy rate is being met. Depending on the institutional comfort level, this process could range from randomly sampling videos for review to routinely checking each video when it is returned.
- Depending on the method of submitting the video to the vendor and receiving it back, the final step of the process may include uploading the video to the desired hosting platform or media player and ensuring that the features all work as intended and are synchronized properly.

Additional workflow steps may be required depending on the specific vendor's approach and the agreement between the parties. For example, in some cases vendors charge by minute, in which case tracking the number of minutes submitted should be included as part of the workflow for budgeting and planning purposes.

Live Event Video Accessibility

Accessibility for live streaming events, particularly those that will be recorded for later distribution as

recordings, is an important workflow to consider when thinking about video accessibility. These steps will help to ensure that both the event and the recording offer maximum accessibility:

- When planning an event, always include accessibility in the plan and the budget from the very beginning. Moreover, it should always be assumed that the event will attract a diverse audience with varied needs; assuming that no one with a particular need will attend is no excuse for excluding an interested participant.
- Select a streaming platform that supports accessibility. An increasing number of platforms have automatically generated captions integrated into the platform, but these suffer from the same accuracy issues as other types of automatic captions. For this reason, it is important to make sure that the platform supports having a stenocaptioner captioning the event as it happens or displaying an ASL interpreter on the screen.
- Ensure that you understand how the platform's features work together. In some cases, captions may be covered by other features, such as chat messages from participants, or the captions themselves may interfere with clearly seeing the ASL interpreter. It is important to check for these issues in advance and, where possible, configure the features and display options to avoid issues.
- Coordinate with anyone who will be speaking or presenting at the event to ensure that they know how to optimize their presentations for accessibility.
- When advertising the event, clearly state which features will be offered, such as live captioning, descriptions, or interpretation, and offer clear instructions for how to request accommodations.
- On the day of the event, have someone available for questions or issues relating to these accessibility features.
- After the event, edit any caption or transcript file for accuracy before posting the recording. Though professional stenocaptioners strive for accuracy, often there will be typographical, spelling, or other errors that need to be addressed to improve the accuracy of the file.
- When posting the recording, post any related files, such as slides that were displayed during the presentation, in an accessible format.

These steps will greatly improve accessibility of the event and the recording and ensure that the content is available to the widest possible audience.

While these workflows may represent new areas of work, they will help to ensure that current and future videos are accessible to users with a range of disabilities. This process is not only legally required in many

jurisdictions, but is also vital to making institutions, their collections, and their programs truly inclusive for disabled users.

Notes

1. World Wide Web Consortium, “Time-based Media: Understanding Guideline 1.2,” in Understanding

WCAG 2.0: A Guide to Understanding and Implementing WCAG 2.0, 2016, <https://www.w3.org/TR/UNDERSTANDING-WCAG20/media-equiv.html>.

2. Laura DeLancey, “Assessing the Accuracy of Vendor-Supplied Accessibility Documentation,” *Library Hi Tech* 33, no. 1 (2015), <https://doi.org/10.1108/LHT-08-2014-0077>.