

January 30, 2019

ITAC Audio Transcription team

Chair: John Grego, Department Chair and Professor, Statistics

Co-chair: Laurel Eddins, Director of Service Operations, DoIT

Team members: Jeff Hostilo, Director of Collaborative Technology & Academic Support, DoIT
Robert Lipe, Clinical Professor, School of Accounting
Michael Kelly, Chief Data Officer, DoIT
Susan Rathbun-Grubb, Assistant Professor, SLIS

Problem statement

The university has a federally mandated obligation to provide accessible learning materials to its student population, including reasonable accommodations for students with disabilities. In addition, the Carolinian Creed obligates faculty to “demonstrate concern for others, their feelings and their need for the conditions which support their work and development.” Unfortunately, the requirements, guidance, and solutions for closed-captioning, transcription, and other accessible learning alternatives available to faculty and staff are unclear. The Faculty and Staff IT Advisory Council plans to research the potential for closed-captioning and/or transcription to solve the following aspects of this problem and make recommendations for future improvement:

1. Which university office(s) should be responsible for ensuring compliance of audiovisual material with mandated accessibility standards?
2. Which office(s) should provide additional guidelines (or best practices) to faculty for accessibility standards of online audiovisual materials?
3. What are closed-captioning and transcription options for teaching and research that USC offices can endorse and support?
4. What available options for closed-captioning and transcription solutions can be readily integrated into faculty workflow?

Investigation methods/data/conclusions

Team members interviewed faculty and staff in the Center for Teaching Excellence, Distributed Learning, DoIT and the Student Disability Resource Center to understand how these offices provided guidance on closed captioning and audio transcription. We scheduled a follow-up conversation with CTE to understand the source for the guidelines they include under their quality assurance standards for distributed learning courses. We relied on the expertise of DoIT staff on our team to investigate enterprise solutions through Ensemble and Blackboard.

We also approached colleagues who teach Distributed Learning to understand what methods they were using. In the Darla Moore School of Business, we talked to staff tasked with video editing to understand their process for closed captioning.

Existing University Resources

Center for Teaching Excellence and Distributed Learning resources

Dr. Aisha Haynes, Associate Director of CTE, confirmed that CTE's Distributed Learning Quality Review program requires both closed-captioning and audio transcripts for video files; over 130 courses have been approved through the program. CTE generally recommends that a transcript could be prepared ahead of recording to facilitate the process. CTE refers faculty to Distributed Learning for transcription services; Distributed Learning provides transcription services through the premium version of Dragon Naturally Speaking. Jeff Soard, Logistics Manager in Distributed Learning indicated that faculty typically make final edits to the transcripts themselves, though Distributed Learning can provide this service as well.

Other alternatives for transcription include YouTube and commercial vendors such as 3PlayMedia, Rev and Cielo 24, though the outsourcing options (the latter three) have become less popular with faculty. Some faculty use Google Voice (Speech to Text) to produce transcripts. USC used to have a contract with 3PlayMedia, though most faculty now prefer to work directly with Distributed Learning. While YouTube is a popular alternative to Dragon Naturally Speaking, the transcript is not punctuated or capitalized, a requirement for CTE's Quality Review program. Distributed Learning found Dragon Naturally Speaking's batch-processing capability to be very helpful. Dragon also has technical libraries that can help with transcription.

Once a transcript is in hand, all contacts recommended uploading the video file and transcript files to YouTube, which will then automatically synchronize the two files to produce closed captioning. It was interesting to learn that this approach was widely adapted by the parties we interviewed.

Dr. Haynes was interested in an enterprise solution, which is used at other universities. She has seen a variety of models, with the work either being done in-house (with student assistance), or outsourced for longer videos. Some universities have budgets for accessible videos and transcription, and some have dedicated offices. Jeff Soard noted that Distributed Learning had attempted an enterprise solution to automate transcription and closed-captioning some years ago using Docsoft, but the product simply did not work as promised.

Distributed Learning would like to make their services more widely available. Currently, most of their business consists of referrals from Center for Teaching Excellence, though their service is advertised on their website. They do not support Palmetto College, but have supported On Your Time and flipped classrooms. The service is free, and will continue to be free.

CTE indicated that faculty must edit transcripts produced by Dragon Naturally Speaking, but Distributed Learning will in fact edit transcripts for small projects and sync the transcript to YouTube. If they were to expand, they could use students to edit transcripts.

Ensemble - Transcription and Captioning Overview

Ensemble is the video platform used as part of the Academic Media Portal (AMP) solution developed by the University of South Carolina Division of Information Technology. AMP is a turn-key solution that features a comprehensive automated workflow using Ensemble Video for the processing, transcoding and publishing of scheduled video capture accessible via BlackBoard or secure web portal. With this solution students can access captured classroom content and

other digital media, specific to each course in which they are enrolled via Blackboard. Through AMP's web portal, Faculty can upload supplemental digital media, generate links to embed content into web pages and maintain an online library of digital assets.

The Ensemble Video platform includes many accessibility features such as integrated speech-to-text technology to automatically create captions for your videos to meet ADA requirements. These automatic captions are generated by machine learning algorithms and include an onboard closed-caption editing tool. Other options such as Interactive Transcript allow you to view the scrolling transcript while watching the video, with current captions highlighted as the video plays. You can also search and click any word to play the video from that point. Ensemble includes multi-language captioning capability where users can create and edit captions in Chinese (Simplified), Chinese (Traditional), French, German, Japanese, Portuguese, Spanish (Mexican), Spanish (Spain), English (US), and English (Great Britain). In addition, Ensemble has partnered with top captioning companies and integrates with 3PlayMedia, CaptionSync, and Rev.

Ensemble's automatic captioning feature is available to its users based on blocks of purchased captioning hours. Hourly rates vary based on amount purchased, as there are discounts for bulk hour purchases. Costs for Ensemble's automatic captioning range from \$3/hr for 300 hours of transcription to \$1.50/hr for 10,000 hours of transcription.

Department of Information Technology

Charity Nix, Certified Microsoft Trainer with DoIT holds a face-to-face 1-hour class entitled: 7 Steps to Make Learning More Accessible. The course covers accessibility features in familiar classroom resources, including Word, PowerPoint, OneNote, etc. It also discusses how to adapt old material and how to create new content with accessibility in mind. Charity also shared Microsoft accessibility links.

Student Disabilities Resource Center

Dow Hammond, Assistive Technology Coordinator at the Student Disability Resource Center, recommends the following services to members of the university community who have audio/video transcription and captioning needs. Turn-around is quick, although it may not always be same-day service.

1. Rev.com - captioning/transcription services are \$1/minute. The first pass is automated and then a second pass is made with manual editing. The results are as accurate as the more expensive services described below.
2. 3PlayMedia.com - \$2.50/minute, but you must upload a minimum of 5 hours of video or have a total of \$750 of captioning needed. Otherwise it is \$3/minute.
3. Cielo24.com - similar to 3Play in pricing.

Dow Hammond referred to a lawsuit against Midlands Tech back in 2016 about captioning (or lack thereof). Their solution was a team of students who serve as captioning assistants using YouTube captioning with hand-editing.

Faculty/Staff Implementation

We talked to faculty and staff who have used transcription services to understand how they typically navigate the process. One of our members talked to a video editor in the Darla Moore School of Business, one talked to an oral historian, and another talked to a convenience sample of four Statistics faculty. Since so many faculty have met the quality review standards, the comments from the Statistics faculty are likely not representative, but do indicate the issues that some faculty face when committing to audio transcription and closed-captioning.

Oral Historian

Andrea L'Hommedieu, Director of Department of Oral History in University Libraries, exclusively uses "real-person" transcription services to transcribe oral histories because of the necessity for accuracy in the historical record. She mentioned that the more accurate, reliable services are national ones. They are more expensive (approximately \$150/hour), but they meet deadlines and are more accurate than the independent transcription services she has used.

She mentioned Dragon as an automated alternative when the person speaking does not change. She also mentioned a new automated online solution called trint.com (<https://trint.com/how-it-works>). Neither of these would work for her transcription needs, but might work for other situations.

Video Editor

Our discussions with Maggie Oswald, resident video editor at Darla Moore School of Business, suggest that closed captions provided by a variety of software packages can achieve approximately 70% accuracy with native speakers, and this percentage can decline dramatically for non-native speakers. When she close captions a video, she does a first pass with Microsoft Stream which is a part of the Office 365 suite and therefore free University wide. Her second pass is to go through the initial closed captions file and manually change inaccurate words. The software used for this phase is Amara, which edits the CC file. To close caption a 3-hour class could easily take 9 hours (more than a full day's work on just that one video) for an experienced video editor; the prospect of closed captioning every video in the Moore School, much less in the entire University, has serious staffing ramifications. A modest improvement to her approach would occur if the University paid for the Ensemble add-on; the first pass in Ensemble is no more accurate, but the second pass is streamlined because the resulting captioned video is already in Ensemble.

Maggie found that creating a written transcript of a closed captioned video is a simple click in most software, while the reverse is not even close to true. Once closed captioned, a subsequent edit of the video will throw off the caption timing and likely require a complete redo of the closed captioning. Thus captioning should be reserved for "final drafts", whether the captioning is done in house or outsourced.

Instructors

One of our team talked to four colleagues who use video in distributed learning classes in both long and shorter formats (e.g., for both lectures and assignment notes). Two use SnagIt or

Camtasia to record videos and simply rely on the automatic closed-captioning produced by YouTube.

Another colleague worked with DoIT, but needed to edit the transcript produced by Dragon Naturally Speaking and has not yet returned to that due to a lack of time and resources; they instead rely on YouTube's transcription.

Another colleague had access to resources not available to most instructors for one course. For the second course, this colleague appears to have taken the same approach others have taken (i.e., default YouTube closed captioning). For the first course, the instructor used YouTube and 3Play Media services to finalize the captions. At the time, USC had contracted with 3 Play Media and the instructor was able to use course lab fees to pay for their captioning service; no transcripts were produced. For a second course that the instructor piloted last summer, the instructor relied on YouTube. Knowing that the class would be taken over by other instructors who would likely replace the videos, the instructor did not go through the expense of paying for a captioning service and editing captions.

Compliance

We made little progress on compliance issues. We talked with Dr. Aisha Haynes, Assistant Director of CTE to discuss how she understands CTE's motivation for developing accessibility standards

Dr. Haynes notes that CTE developed accessibility standards as part of the Provost's 2013 call for Distributed Learning Quality Review (the Quality Assurance rubric includes 49 different standards). CTE worked with Student Disabilities Resource Center (SDRC) to develop access guidelines. Dr. Haynes indicated that CTE took their cue from SDRC, using their best practices to develop six accessibility Quality Assurance standards. Quality Matters, an online standards portal, was the primary source for standards, though their guidelines can often be vague and need to be filled out (e.g., "provide accessible documents" is converted to directives like "require OCR on all scanned pdf's"). Dr. Haynes felt the standards were based on an interpretation of national laws/guidelines for accessibility.

Incidental to our study of the compliance issues, the committee learned that the Office of Civil Rights (OCR), a federal agency, selected USC and other universities to investigate website accessibility under the Rehabilitation Act of 1973. The Office of the President convened a special website accessibility committee headed by members from Office of Equal Opportunity Programs and Student Disabilities Resource Center, who arranged OCR training for all unit website managers. It was through the scheduled training that we learned of the planned investigation.

It seems that the public website is not the sole focus of OCR. They are apparently also looking at online learning materials which are restricted to enrolled students. Depending on the outcome of future meetings and decisions, the closed captioning of videos intended for student learning may become a high priority. However, OCR's presentation mentions that entities "are not required to make adjustments ... that ... impose an undue financial or administrative burden on the entity." Based on what the committee has heard, captioning "one-off" videos that are

used in a particular semester and then unlikely to ever be used again when no registered student needs such an accommodation may meet the criteria for “burdensome”.

Executive Summary

The university has reasonably-priced solutions to audio transcription and closed captioning services, though implementation for an entire course will always require a significant amount of planning and time on the part of the instructor. Distributed Learning already has on hand much of the resources to meet much of the demand for closed captioning and audio transcription, and additional support for their services comes at minimal cost. Regardless, enterprise solutions and some commercial solutions are appealing too, provided a funding mechanism for accessible online resources can be identified.

Decisions about compliance appear to originate from individual offices’ interpretation of federal law. We did not identify any clear structure or plan for translating federal regulation into USC policy, nor identify authority within the administration for enforcement of guidelines.

Recommendations/Actions

It is clear from discussions with CTE that audio transcription and closed captions are only a small part of the accessibility equation and general quality review for online courses. The CTE is an excellent resource for advice and assistance on course development, and for certifying course design.

The committee recommends that further talks be conducted with Distributed Learning to understand the volume of audio transcription and close-captioned services they could provide gratis. In some respects, their offer sounds ideal (too good to be true?) and their model—transcription by Dragon Naturally Speaking, editing by the faculty or DL staff, then syncing via YouTube—works well. Distributed Learning feels they could scale their model relying on student help, so costs would remain low. Otherwise, commercial and enterprise solutions are substantial for single course costs, apparently costing more than \$100 per lecture (\$2800 per course), and perhaps a commensurate amount for supplementary video materials. The committee did not explore funding, though funding from the departments themselves may be possible provided that they receive a reasonable share of distributed learning course revenue.

Some faculty and units navigate the process of transcription and closed captioning efficiently, while others find the process burdensome, and the tools at their disposal are often inefficient. Some of these differences are likely due to differences in course delivery, depending on whether recordings are used only once or used many times before being replaced. The university should support an enterprise solution, and university offices should better advertise services available for faculty.

Currently, the Academic Media Portal solution, including the Ensemble video platform, is only funded as a solution for Distributed Learning courses supported by the Division of Information Technology. As an academic support organization, DoIT would like to lead a university dialog concerning the potential enterprise adoption of AMP and the continuous development of a roadmap for lecture capture, accessibility and digital media management.

Based on conversation with CTE and our knowledge of the ongoing OCR investigation, it appears that university offices have a fair amount of autonomy to develop standards, and the motivation to develop standards is situational. The President's formation of an ad hoc website accessibility committee suggests the need for a more formal structure involving the Faculty Senate, Office of Equal Opportunity Programs, the Office of the Provost and CTE, DoIT, and the Student Disability Resource Center to provide guidance on accessibility standards.

Resource requirements and strategies

With captioning costs in the range of \$1/min to \$2.50/min, resources to support a course would require a reasonable investment. Solutions that require intensive faculty involvement will need support from DoIT, CTE and the Student Disability Resource Center.

The use of Ensemble as an enterprise solution would require more than the limited funding currently provided for DoIT supported Distributed Learning courses. Cost would be directly impacted by the scale of users and amount content to be accommodated. Fortunately, Ensemble's diverse architecture allows for several deployment options that are scalable and easily adjusted as the university's needs change. Currently, our deployment is a self-hosted model with an annual recurring user license based on 40,000 maximum full time enrollments. Along with the user license costs, we have an additional expense of several on-premises virtual servers for encoding, transcoding and hosting the web based application and additional on-premises storage for all digital content.

Determining the true cost of an enterprise deployment of Ensemble has to start with defining the following considerations, many of which are beyond the scope of this committee's charge.

1. Needed enterprise user license size. Cost saving can be found in multi-year licensing.
2. Needed hosting model, or growth plan to transition model as we expand (e.g. self-hosted, cloud-hosted, hybrid hosting).
3. Virtual hardware needed to accommodate encoding, transcoding and hosting the web based applications. (Growth plan will define milestone growth point which will dictate the need for additional hardware.)
4. Features and Functionality needed (e.g. captioning, screen capture, live streaming, etc.)
5. Content Security and Governance
6. Storage
 1. A digital asset management strategy will help determine the amount of needed active and archive storage.
 2. Costs will be dictated by on-premises and/or cloud storage rates.

DoIT recommends that an advisory committee be formed to research and advise the administration regarding an Ensemble enterprise digital content storage model and determine retention policies, governance and associated costs. DoIT, CTE, ODL, Faculty IT Committee and representatives of each academic unit should be included.

Appendix I: Useful weblinks

Quality Matters: <https://www.qualitymatters.org/>

Guidance for YouTube captioning service:

<https://support.google.com/youtube/answer/2734796?hl=en>

Dragon Naturally Speaking: <https://www.nuance.com/dragon.html>

Trint: <https://trint.com/>

Microsoft resources

Microsoft's main accessibility site: <https://www.microsoft.com/en-us/accessibility>

Windows 10 Accessibility: https://www.microsoft.com/en-us/accessibility/windows?activetab=pivot_1%3aprimar2

Office 365 accessibility: https://www.microsoft.com/en-us/accessibility/office?activetab=pivot_1%3aprimar2

Disability Answer Desk: <https://www.microsoft.com/en-us/Accessibility/disability-answer-desk>

Blog post real educator story - <https://blogs.office.com/2016/10/19/accessibility-in-the-classroom-tools-that-impact-my-students/>

Using more accessible styles and colors - <https://support.office.com/en-us/article/video-use-more-accessible-colors-and-styles-in-slides-65c8c964-9449-4b5f-88ea-e226abd68337?ui=en-US&rs=en-US&ad=US>

Accessibility Sway - <https://sway.office.com/nZixyVhU9cHYdUS0>

Free training course in the Microsoft Educator Community -

<https://education.microsoft.com/courses-and-resources/courses/training-teachers-to-author-accessible-content>