Vanessa Castaldo, Emily Caspari, Laura Indick, Simi Best
INFO 665: Projects in Digital Archives
May 7, 2020

Report: Digitization Team

We were tasked with digitizing two analog collections. The first was an audio reel-to-reel collection from the Lesbian Herstory Archive, of the radio program "Lesbian Nation." The second was a video collection on U-matic tapes from the Lesbian, Gay, Bisexual, & Transgender Community Center of New York, of the local television program "Gay Morning America." We used an UltraStudio 4K Black Magic digital converter and Black Magic Media Express software for digitization. To convert to access copies, we used Adobe Media Converter. We used Audacity Software to convert the audio reel-to-reel recordings into digital files.

Work Completed:

For the audio digitization project, we completed six reel-to-reel tapes, and four tapes were partially completed but broke during the process. For the video digitization project, we completed nine U-matic tapes. We successfully made master and access copies for each of the completed digitized tapes.

Challenges:

The very first time we tried to digitize a file, Audacity crashed, and although the file seemed like it was recovered when Audacity reopened, the file was corrupted, with random gaps of silence in the recording. This made us realize that we had to keep a fairly close eye on the image of the sound waves, to note if anything looked strange, and gave us a sense of the various things we had to check to make sure that a file had been digitized completely.

The second day (February 13), we had some new technical issues. We came before class, hoping to spend 2 hours digitizing the audio, but we weren't able to do any, because, although the open reel player was working, the connection to the computer was not working, because Audacity was not recording any audio. (We also struggled a bit to get the tape to stay in the slot, although we eventually were able to figure it out). We ended up waiting for Professor Cocciolo to come in, because we didn't want to mess anything up by trying to fix it. Later the same day, once we had the system working, the first audio tape that we connected snapped almost

immediately, with the first two inches or so of tape breaking cleanly from the rest of the reel. Although we definitely wanted to try splicing the tape back together, we took Professor Cocciolo's advice and put them aside into a "needs special attention" section. We put a note, physically (via post-it and piece of paper) and digitally (in our inventory spreadsheet) noting that the tape needed special attention. We decided that we would put these tapes with special issues aside until the end of the semester, and our decisions about attempting actual interventions with them would depend on the amount of time we had left when we got to them. Professor Cocciolo looked into it and found a specialty website that sold open reel splicing tape and leader reel, and we decided to go ahead and order both so that we'd have them in case we ended up using them.

As February 13 was our first day trying to digitize the video, there was a little bit of a learning curve for that as well, although in general it was much easier than the audio files. One thing that we learned was that we needed to set a timer for 58 (or 30, depending on the season) minutes when digitizing the "Gay Morning America" tapes. For the first tape, we didn't yet know how to do this, and ended up with about 2 minutes of just a black screen at the end of the file. We decided to wait until we had exported the file to try cutting out that space. In the meantime, though, we added that to the workflow.

On February 20, we had a number of technical issues. An audio tape snapped, and as we did not yet have our splicing equipment, we weren't able to do anything about it. The break was across the intake reel, the tape reel, and two pieces of tape, which we put aside in our "needs conservation" area. With video, the first digitization worked well, but we struggled to save the access file correctly, as the settings were different in Adobe Media Encoder. However, we were able to figure it out (with help from Professor Cocciolo), and edited our workflow document to reflect the necessary changes. The second videotape that we digitized had image issues -- the middle of the tape (about 10-15 minutes) had major image distortions, although the audio was mostly intact. We were at first worried that the U-matic tape player was failing, but as the remainder of the tape played mostly normally, we were relieved to realize that it was an issue with the tape itself.

On February 27th, as the audio tape was digitizing, it was obvious that tape 008 had been heavily spliced. The audio sounded fine as it was playing, but it was visible that pieces had been spliced. One of the little bumpers on the player would slightly pop up everytime a splice would pass through, so this tape needed quite a bit of attentive supervision. Once it got to the 25 minute mark, it snapped at yet another splice point. The box indicated the interview was 30

minutes, so we kept the incomplete recording until we decided as a group and with Professor Cocciolo whether we want to splice it and risk it breaking upon a second recording. The worst part about this was that we'd lost yet another take-up reel until we were able to splice and reload it back onto one reel. There was an extra take-up reel in the classroom, but it was the last spare, so we hoped we would not experience another breakage. We experienced another breakage this same day, this time with tape 010, which was backwards and needed to be fast-forwarded onto our take-up reel so we could play it back in the right direction. As the tape was mostly loaded onto the other reel it snapped at yet another splice point. We realized that we were going to need to splice some of these tapes because we were out of spare reels, as all three were occupied with portions of other broken reels. Another challenge this day was that only Vanessa was able to come digitize outside of class time that day, and with only one person it was not possible to digitize both formats at once.

On March 5, we had an issue beginning the video digitization, as Black Magic Media Express wasn't showing the usual dropdown menu for input & output options. With the help of Professor Cocciolo, we were able to get it working. However, we weren't able to do the open reel digitization, because although the extra re-uptake reels had arrived from eBay, it seemed that they were the wrong size reels. We discussed the possibility of trying to splice the open reel tapes that have broken, but we decided to wait until we have the correct size uptake reels to try anything.

On March 9, we had the same issue beginning the video digitization, in which the Black Magic Media Express was not recognizing that a device was connected, and so we were unable to change the input and output options in the regular dropdown menu. Unplugging and plugging back in the Blackmagic machine was not working. We eventually figured it out by crawling around the floors and tables to trace the paths of two Thunderbolt cables that were connected to the Blackmagic machine, and we discovered that one of the Thunderbolts was connected to the computer to the left of the one we normally use. Once we plugged both Thunderbolts into the correct computer, the device was connected, so the cables must have gotten switched somehow. We decided to make sure to plug both in before digitizing just in case.

That day, we also noticed that one of the access copies (gca_gma_02jan1984_access) we had made in a previous week was not playable in quicktime and was five times larger than other access copies, despite the fact that the videos were similar runtimes. We also noticed that this video, and another previously digitized video (gca_gma_26dec1984_access), were not playable directly in Google Drive's media player the way all of the other access copies are. We decided

to start over converting the master copies of both videos to the access copies. It worked when we started fresh, and we don't know what the problem was.

Soon after March 9, we faced our largest challenge by far: our project was cut short by the COVID-19 pandemic, which shut down Pratt Manhattan and forced us to all isolate separately. At that point, we weren't able to digitize any further.

Discussion/Decisions

For the audio tapes, we decided to create a "Done" box, where we would put the tapes once they had been digitized. We labeled the original box, with all the tapes "To Do," so we could know right away that any tape we took from there had not yet been digitized. It was a little more difficult for the video, because some of the tapes had been digitized by the previous semester's class. On the third working day (Feb. 13), we decided to go through the boxes of video tapes and divide them into two categories: already digitized and to be digitized, to cut out a step from the workflow.

For the video files, because the 10-bit YUV files would be about 100 GB per hour, we knew that we would definitely need to create smaller access files to upload to the site. We decided to follow the lead of the previous semester's group, and use the Adobe Media Encoder to export as H26 MPEG4 files, choosing the "High Quality 450p SD" option for the access files. Then we uploaded these to the Google Drive, using their same file naming format as well, to create the most continuity possible for the project. We also used the same hard drive to save the larger master files too, based on the lesson learned during the previous semester about trying to save those to Google Drive. We were glad to be able to learn from their experience!

For the audio files, we decided that we didn't need to create separate master & access files, because the files weren't large enough to pose a problem. We created a system for keeping an inventory, saving the files to the Google Drive, and recording our progress as we went along. We tweaked some small things in our system as we went along (for example, we learned that we couldn't hyperlink the two images in one column into a Google Sheet, so we added an extra "photo" column.)

For inventory of the audio tapes, we talked with the metadata team, because we were concerned about doing double data entry unnecessarily, and we decided that they would do a full inventory of both sets of tapes, allowing them to control the schema that they decided to

use, whereas our group would just assign numbers and take photos with the number in them. However, we got a bit nervous about having some sort of content marker for the audio tapes so, if something went wrong, we could identify the file by listening to it, so we added a "Keyword" column in the inventory. For the first one, for example, we decided on keywords "Interview", "Frances Doughty", and "Sappho," so if someone came across the audio file and for some reason couldn't identify it based on the file name or related photo, they could listen to the intro, which states that it is an interview with Frances Doughty about Sappho, and know which tape it was from. At first we took photos of both the front and back of the tape boxes, but by the third working day (February 13), we had decided to simplify the system and only take a photo of the side that had writing on it. If there was no side with writing on it, we would still take one photo for identification purposes. For the file naming format, we wanted to keep it simple, knowing that the file names were only internal and would not be viewed by the public, so we decided to only include the title of the organization (LHA), the format (Audio), and the number that we had assigned the tape for the purposes of digitization, which we knew the metadata team would likely change

In terms of deciding what to digitize, we picked the materials that seemed like they would be most easily digitized, especially once we began having problems with the reel-to-reel tapes. So, for the audio tapes, we prioritized the tapes that had intact leader tapes and that looked like they had not been previously spliced. For the video tapes, we prioritized the full episodes over the other content.

Future Directions for the Project

It's unfortunate that we weren't able to finish the semester, and because our digitization efforts were cut short unexpectedly, future students will have to pick up from where we left off. Because we didn't get a chance to splice the reels that broke, future students will have to work on splicing the broken reel-to-reel tapes and fixing the two Gay Morning America access copies that don't work. This taught us to take into account the "hit by a bus" rule of archival work -- to be sure to leave things every day in a place where it's possible for someone completely different to pick up the project at a later date. In a way, this was very realistic to "real-life" archival work!