

## DIY VHS Preservation: Planning for video digitization at the American University Library

### **Abstract:**

Most videotape preservation specialists agree that there is a discrete window of time, an estimated 10-12 years, for owners of videotape collections to undertake and complete digitization projects to save their out-of-release content for future use before forces converge to make those efforts all but impossible. The combination of obsolescence of playback equipment, lack of replacement parts, degradation of magnetic tape, and loss of repair expertise will make large-scale digitization projects technically and economically unfeasible.<sup>1</sup> At the American University (AU) Library, 2364 of our commercially-distributed videotapes were identified as being out-of-release. An additional 500+ hours of university-produced content needed to be preserved. Due to pressing needs to protect high-use, but deteriorating, VHS and U-matic tapes and the prohibitive costs of out-sourcing digitization, the Media Librarian and the Visual Media Collections Coordinator undertook building an in-house digitization rack with attention paid to long-term archival quality and viability. This article describes their planning process.

### **The Situation:**

American University is a medium-sized liberal arts university in Washington, DC. With 11000 FTE students and 725 FT faculty, the curriculum and library collections span most of the social sciences, humanities, physical sciences, and business. Across the curriculum, AU faculty members have been active users of visual media to complement their classroom teaching. Accordingly, the AU Library's visual media collections were built to address this need. The range of content is broad and includes video recordings from educational distributors, mainstream retailers, and work distributed by filmmakers and producers directly. The AU Library also has a wealth of campus recordings including commencements, guest speeches, and sporting events. To be exact, 2364 of our commercially-distributed videotapes have been identified as being out-of-release and an additional 500+ hours of university-produced content needs to be preserved. Each of these collections includes substantial content that is irreplaceable and at risk of loss yet still of considerable value for instructional, research, and historical interest.

### **Earlier AU Library media preservation efforts:**

In the mid-1990s, prior to the easy availability of video digitizing hardware and software, the AU Library's Media Services department had a videotape crisis. The department's collection of approximately 1000 U-matic videotapes fell victim to rapid deterioration and consequent failure of the U-matic tapes and playback equipment. Tape deterioration due to age and environmental factors contributed to equipment breakdowns. It was apparent that many U-matic recordings, from distributors such as PBS and Films Inc. were no longer available in any format. At the time, permission to transfer the tapes to VHS was obtained from several distributors and deteriorating, out-of-release content from shuttered distributors was transferred under the guidelines of Section 108 of the U.S. Copyright law. VHS preservation recordings were made with a tape-to-tape process using analog video recorders and with minimal quality control measures.

In the late 1990s, when the DVD format was more widely adopted, there was considerable effort by libraries to duplicate VHS content in the new format. A steady stream of re-released content required that budgets be divided between purchasing new titles and replacing high-use older VHS content. At the time there was little thought and no funding set aside for preserving out-of-release VHS materials. Over the past fifteen years, the flow of VHS content re-released on DVD has gradually ebbed to a trickle and now it is apparent that a substantial portion of those VHS collections are out of release in any format. 28% of the total VHS holdings at AU are out of release.

### **Current Digitization Station and Workflow at American University**

<sup>1</sup> See Mike Casey's article, "Why Media Preservation Can't Wait: the Gathering Storm" referenced as links 1 and 2 under 'Additional Resources' at the end of this article.

## DIY VHS Preservation

The following explains the workflow developed by the AU Library Media Services unit to digitize at-risk commercial and campus-produced content for preservation. Thorough research into digitization technology and preservation standards was conducted in the development of the process; however, our method is not the only viable option for conducting digitization work. Among other factors, cost efficiency could not be ignored. The cost of creating and storing uncompressed video files of all our tapes could not be justified with our budgets. This workflow represents what was determined to be the best feasible procedure for the AU Library VHS collections at this time.

What follows is a summary of the VHS digitization preservation process including planning, identification of highest-risk materials, hardware, software, storage of digitized files, and documentation. With ever-changing technological advances, the Visual Media Collections Coordinator will continue to update the practices followed. The description focuses on the most cost-effective and viable workflow for the collections at hand. A list of online resources with information that could inform alternate workflows is appended.

### **Identification and evaluation of highest-risk materials**

The most pressing problems are related to the commercially-distributed circulating collection. Circulation reports for this collection have been used to pinpoint the oldest titles as well as those with highest and most recent usage. In each case, the title is searched in Worldcat, Amazon, and in a general internet search for evidence that it is currently available for acquisition on DVD or Bluray. When replacements can be acquired, they are. When they can't, the searches are logged using a Copyright Search Log<sup>2</sup> created by the Media Librarian and the title is added to the queue of items to be digitized. This selection process adheres to the guidelines spelled out in Section 108 in the U.S. Copyright Law, permitting university libraries to preserve at-risk items when unused replacements aren't available.

The other major component of the VHS preservation project is campus-produced content. The AU Library has collected hundreds of VHS recordings of lectures, readings, athletics events, commencements, and other university events on analog tape formats. Most of these recordings have minimal metadata and are unique. It is a far more time-intensive project to analyze and evaluate this university-produced content than the commercially-distributed content. The identification and evaluation of these collections involves working from the minimal tape labels and any box lists that come with the assets and then investigating dozens of unlabeled or poorly labeled items to determine what is on each tape and its value to the Archives. The Visual Media Collections Coordinator works with the University Archivist to decide which items the Archives will officially add to their collection. When selected for preservation, campus recordings are digitized at the highest possible standards available to us.

### ***Building a digitization workstation:***

Even if the highest quality preservation files may not be currently feasible for the thousands of video recordings in need of preservation due to lack of storage, it is crucial to have the proper equipment to digitize items and create archival-quality files when needed. There are a variety of products available to achieve this. Listed below are the hardware and software choices the department made during the 2013/2014 academic year to create the video digitization workstation, all of which are still being used:

#### Hardware:

1. 2013 iMac 27" with 1TB HD space, 16GB RAM, and a 3.4 GHz Processor
2. Blackmagic Intensity Shuttle for Thunderbolt™
3. AVToolbox AVT-8710 Multi-Standard Time Base Corrector

2 See Copyright Search Log referenced as Link 3 under 'Additional Resources' at the end of this article.

## DIY VHS Preservation

4. Analog video players (VHS/S-VHS/miniDV/Betacam/SuperBetamax/Laserdisc/U-matic)
5. Promise Technology 32TB Pegasus2 R8 Thunderbolt 2 RAID storage array (7200 rpm)
6. TV for monitoring capture process, plugged into Blackmagic Intensity Shuttle

### Software:

1. Blackmagic Media Express for video capture
2. Adobe Premiere Pro and Adobe Audition CC for editing and finalizing files
3. Adobe Media Encoder CC for creating mezzanine and access versions
4. Library of Congress Bagger transfer tool and HashX (from BoilingBit Software) for creating and verifying checksums
5. Microsoft Excel for metadata management
6. Roxio Toast Titanium for creating preservation DVD access copies

### **Basic Steps for preserving VHS recordings:**

Outlined below are the steps AU follows to ensure creation of the highest quality digital files from the VHS recordings. While ideally, the master preservation files would be captured in 10-bit uncompressed video codec, due to the scale of the project and storage constraints, the master preservation files are currently being captured in uncompressed 8-bit codec. To illustrate what this means: the size of a two-hour recording saved as an 8-bit uncompressed MOV file is approximately 150 GB and a 10-bit uncompressed file of the same recording would be approximately 200 GB.

### **File Creation**

For all incoming AU-produced content, once an item is selected for preservation and scheduled for digitization, the following steps are taken to create digital files and prepare them for storage. The process was adjusted for the commercially-distributed content. These changes are explained below.

1. Digitize the tape using an analog video player connected to the iMac via the Blackmagic Intensity Shuttle for Thunderbolt™ and time base corrector using the Blackmagic Media Express software. Recordings are saved as 8-bit uncompressed MOV files with Linear PCM audio codec. The output is the draft master preservation file.
2. Using Adobe Premiere Pro CC, the digitized video is trimmed to the program length and the image resized to fill the frame. Also any basic audio and video touch-ups are done in Premiere Pro and Audition, such as removing hum or simple color correction. This finalized master preservation file will ideally be stored in deep preservation storage and monitored regularly for file integrity. Otherwise, it goes untouched unless all other files are lost or damaged.
3. Using Adobe Media Encoder CC, the master preservation file is transcoded to an MOV with H.264 video codec and Linear PCM audio codec. This is the mezzanine file AKA production master file. This file serves as the quick-access master copy and back-up in case the access file is lost or damaged. Having this file allows you to keep your master preservation files in deep storage undisturbed while still maintaining access to a decent quality version of your asset.
4. Using Adobe Media Encoder CC, the master preservation file is transcoded to an MP4 file with H.264 video and MP3 audio codecs. The output of this process is the access file, used for streaming and primary viewing.

Version	File Format	Video Codec	Audio Codec
Preservation Master	.mov	Uncompressed 10-bit/ Uncompressed 8-bit	Linear PCM or .bww

## DIY VHS Preservation

Mezzanine (Production Master)	.mov	H.264	Linear PCM or AIFF or .wav
Access File	.mp4	H.264	.mp3 or AAC

5. Once all copies are created, they are saved together in a folder with a unique file name. In a Microsoft Excel spreadsheet, the digital file names are added to the asset metadata.
6. Using Library of Congress' Bagger tool, the three tiers are "bagged" into a folder and checksums are run on all three files. The checksums are automatically stored in the folder with the three files but they are also added to the spreadsheet. Checksums will be run annually to verify the integrity of the files stored locally and on network drives.
7. An access DVD copy of the program is made using the mezzanine file for either the Media Services collection to house for on-site viewing or for University Archives to maintain for research purposes. Roxio Toast Titanium is used to auto-compress the file to fit on an archival-quality DVD (such as Verbatim UltraLife Gold) to be used as a playable disc by researchers and/or instructors.

As previously mentioned, this process is being followed for the campus-produced content but not for the out-of-release commercially-distributed content. Due to storage and time constraints, the Visual Media Collections Coordinator adjusted some of the steps in the process for VHS content in the library's circulating collection. The commercially-distributed content is captured in DVCPro50 video codec instead of uncompressed 8-bit. This produces a 52 GB file instead of a 150 GB file for a two-hour program. The Linear PCM audio codec remains the same, as does the hardware and software used to capture the files.

Furthermore, the mezzanine and access copies for the commercially-distributed content are not yet being created so more tapes can be digitized before the tapes and equipment fail beyond repair. The goal is to return to the digital inventory and create the derivatives to ensure proper preservation standards are met.

Since there is only one digital file per title to save for this collection, checksums for the commercially-distributed content are run using HashX instead of Bagger. All other processes, including file clean up and access DVD creation, are the same. Once storage issues are resolved, the preservation process will become identical to that of the campus-produced content.

### Metadata

Basic program and technical metadata is saved in the previously mentioned Excel spreadsheets. The campus-produced content that the AU Library receives rarely comes with metadata, so the information about the tapes is minimal and primarily consists of what is written on their labels. Copyright questions, checksums, dates digitized, digital file names and location are added to the metadata spreadsheet. As these recordings are analyzed more closely, additional metadata will be collected and created.

Program metadata about the commercially-distributed VHS recordings exists in the library's online public-access catalog, so the metadata spreadsheet for this collection contains minimal program metadata. Basic technical metadata of the digital file is maintained in the spreadsheet.

Caption tracks are one metadata element of note. We are capturing any existing line 23 caption tracks during digitization. However, we currently do not have software that can read and extract those captions to display on our digitized files or access DVDs. We are investigating possible solutions to extract captions as part of our regular workflow. But for now we are notating which digitized items contain the captured line 23 caption track and we are working with our Accessibility office on caption requests on an as-needed basis.

### File Storage

## DIY VHS Preservation

At present, most digitized files are stored on two separate Promise Technology 32TB Pegasus2 R8 Thunderbird 2 RAID storage arrays housed in the library – one for commercially-distributed content and one for archival content. When a folder or file is moved to a RAID array, the checksum is validated to ensure the file is unchanged during transfer. More stable, long-term storage solutions are being investigated now.

A small selection of commercially-distributed tape files were sent to the Washington Research Library Consortium (WRLC) Storage Facility, of which AU is a member, to be backed up on a networked server where they can be monitored with fixity checks and to ensure a duplicate copy is stored at a separate site from the Library. This could be a viable permanent solution for the rest of the collection, with workflow adjustments and increased storage. Since the files are so large, uploading/downloading to an online solution is not feasible, which leaves transferring hard drives back and forth between the library and the off-site storage facility. This is a fine, albeit unwieldy, solution that requires a lot of time and needs concrete understanding and seamless communication between facilities. Additional off-site storage options, including LTO tapes and cloud storage for smaller files, are under consideration in addition to the network storage at the off-site storage facility.

### Time commitment estimates per tape for three-tier highest quality preservation:

Step	Duration (approximate)
Digitization from Analog to Digital	Real time (length of tape)
Minor clean-up of file in Adobe Premiere CC	0.25 hour
Transcoding Uncompressed to Mezzanine file & Access copy	0.5 hours
Running checksums on all tiers	1 hour
Burning a DVD access copy	0.50 hour
Copying folder to RAID array (auto-back up)	0.50 hour
Verifying checksums from RAID array	1 hour
Technical Metadata entry	0.25 hour
<b>TOTAL TIME:</b>	<b>6 hours for a 2 hr. program</b>

### Conclusion

When tackling the question of where to begin when digitizing a large circulating video collection, the first priority for the AU Library was to protect high-use videotapes showing obvious signs of wear. It was then necessary to complete a comprehensive inventory of the distribution status of each item in the collection before the next stages of the project could begin. In the spring of 2016, the Media Librarian and Visual Media Collections Coordinator completed a year-long comprehensive inventory of the copyright status and availability of AU's circulating VHS collection of 8005 items. With the knowledge gained from this research, all videotapes that were already replaced on DVD or are currently commercially available – over 5000 titles in all – were sent to an off-site storage facility. The material that remained, 28% of the total collection or 2364 tapes, were the VHS tapes that needed to be digitized.

The next stages of the digitization project included the prioritization of the remaining tapes and determination of a timeline for its completion. With this established workflow and equipment set up, the Visual Media Collections Coordinator, working with part-time student staff, was able to simultaneously address the preservation of both the circulating collection and campus-produced video. To date, the Visual Media Collections Coordinator allots approximately 30% of her weekly schedule to the digitization and preservation of these recordings, at a pace of approximately 300 videotapes annually, focusing on

## DIY VHS Preservation

circulation statistics and rarity. Depending on budget and time constraints, triage and de-accessioning of some content may also be necessary.

Of the 300 videotapes digitized annually, approximately 100 are campus-produced content. In addition to the three-tier preservation files and DVD copies, much of this archival collection will be uploaded to the American University Digital Research Archive ([aislandora.wrlc.org](http://aislandora.wrlc.org)), an open-access repository with streaming capabilities. This collection will include commencement speeches, athletic events, special events, guest speakers, and student TV productions.

We believe the process developed at the AU Library is cost-effective and sustainable for medium to large university libraries with circulating analog video collections and a backlog of campus recordings. Given the increasing number of ad hoc requests the AU Library has received from faculty members and departments seeking to save video recordings, we recommend that every academic library consider building a digitization workstation and have personnel trained to use it and assist others with videotape preservation projects. The AU example shows that the expertise needed to assemble a workstation and undertake a preservation project can be attained with a reasonable effort by a capable employee in the course of their regular duties. To delay this kind of project for even a few years may mean that a substantial amount of irreplaceable content will be unnecessarily lost.

### **Additional Resources:**

- 1) Slideshow of Mike Casey's AMIA Presentation on Degradation and the Crisis of AV Digitization: [http://www.avpreserve.com/wp-content/uploads/2013/11/casey\\_amia2013.pdf](http://www.avpreserve.com/wp-content/uploads/2013/11/casey_amia2013.pdf)
- 2) Mike Casey's article, "Why Media Preservation Can't Wait: the Gathering Storm" (IASA members only): <http://www.iasa-web.org/iasa-journal/iasa-journal-no-44-january-2015>
- 3) Copyright Search Log : <http://aislandora.wrlc.org/islandora/object/mediaservices%3A4>
- 4) WRLC Digital Project Guidelines and Information: <https://www.libraries.wrlc.org/dpca/proj-info/Digital%2bProject%2bGuidelines%2band%2bInformation.html>
- 5) Library of Congress Digital Preservation page: <http://www.digitalpreservation.gov/>
- 6) Library of Congress Information on Digital Formats: <http://www.digitalpreservation.gov/formats/>
- 7) AMIA Videotape Preservation Fact Sheet: [http://www.amianet.org/sites/all/files/fact\\_sheets.pdf](http://www.amianet.org/sites/all/files/fact_sheets.pdf)
- 8) AMIA Storage Standards and Guidelines: [http://www.amianet.org/sites/all/files/storage\\_standards.pdf](http://www.amianet.org/sites/all/files/storage_standards.pdf)
- 9) Digitizing Video for Long-Term Preservation: An RFP Guide and Template (NYU Libraries): <https://library.nyu.edu/preservation/VARRFP.pdf>
- 10) NYU MIAP VIDEO AT RISK: Strategies for Preserving Commercial Video Collections in Research Libraries PDF: [http://www.nyu.edu/tisch/preservation/research/video-risk/VideoAtRisk\\_SECTION108\\_Guidelines\\_2013.pdf](http://www.nyu.edu/tisch/preservation/research/video-risk/VideoAtRisk_SECTION108_Guidelines_2013.pdf)
- 11) Bay Area Video Coalition: <https://bavc.org/preserve-media>
- 12) Smithsonian Institution Born-Digital Video Preservation: [https://siarchives.si.edu/sites/default/files/pdfs/bornDigitalVideoPreservation2011\\_0.pdf](https://siarchives.si.edu/sites/default/files/pdfs/bornDigitalVideoPreservation2011_0.pdf)
- 13) Activists' Guide to Archiving Video: <http://archiveguide.witness.org/>
- 14) Multimedia Wiki: [http://wiki.multimedia.cx/index.php?title=Main\\_Page](http://wiki.multimedia.cx/index.php?title=Main_Page)
- 15) JISC Digital Media Digitization Page: <http://www.jiscdigitalmedia.ac.uk/digitisation>