Welcome to Audiovisual Alacrity: Managing Timely Access to Audiovisual Collections

Siobhan Hagan graduated from NYU’s Moving Image Archiving and Preservation program and began her career at the UCLA Library in 2011 as their first Audiovisual Preservation Specialist. In January of 2014, Siobhan began working at the University of Baltimore as the first Audiovisual Archivist in the Langsdale Library Special Collections Department. Her passion is the preservation and access of regional audiovisual materials and collections, as well as advocating for the immediate reformatting of magnetic tape as the world nears a catastrophic loss of recordings stored on these ephemeral and obsolete formats.

Steven Villereal is the Audiovisual Conservator in the Preservation Services department at the University of Virginia Library, where he oversees the reformatting of legacy AV materials as well as consulting on the preservation and curation of born-digital production projects. He is a graduate of New York University’s Moving Image Archiving and Preservation program.

I am Robin Pike. I graduated in 2007 from the University of Pittsburgh with an MLIS, specializing in Archives and Records Management. After working for four years as the Audiovisual Archivist at The Catholic University of America in Washington, DC, I became the Manager of the Digital Conversion and Media Reformatting Department at the University of Maryland Libraries in February 2012. I am the current chair of the SAA Recorded Sound Roundtable, and serve on the Education and Training Committee of the Association for Recorded Sound Collections. My goal is to promote and educate archivists, librarians, and other information professionals on the timely preservation of audiovisual media.
Our recorded history is at risk.

In all our repositories, boxes are filled with jumbled cassette tapes mixed with sticky shed open reel tapes and 8-track cartridges. VHS and U-Matic tapes stacked on open shelves silently degrade. In the farthest corner of our shelving, away from any other materials, sit di-acetate films that reek with vinegar syndrome.
Large-scale studies by leading institutions like Indiana University have confirmed our worst nightmares—we have about 10-15 more years to preserve what we really want to preserve, and the media and playback equipment might not even last that long.
As Mike Casey says, Degralescence, the two-headed monster of degradation and obsolescence, lurks in our archives.

Preserving media is a problem, both financially, and because most of the time the containers don’t have enough descriptive metadata for us to gauge if the content is valuable enough to justify the cost of preservation and digitization. We will not be able to save all of the history we’ve been tasked to protect but we need to act now because audiovisual preservation is a time-consuming process.

The three of us will discuss different sustainable models for managing the timely description, preservation, and digitization of audiovisual material, within the collection department, the preservation department, and the digitization department.
Audiovisual Alacrity: Managing Timely Access to Audiovisual Collections

Siobhan Hagan, University of Baltimore, Langsdale Library
Robin C. Pike, University of Maryland Libraries
Steven Viltereal, University of Virginia Library
My name is Siobhan Hagan and I am the first Audiovisual Archivist at the University of Baltimore’s Langsdale Library Special Collections Department. The University of Baltimore is one of twelve institutions in the University of Maryland System and is located in Baltimore City right next to Pennsylvania Station. The majority of our AV materials lies in the collections of two Baltimore-area television stations: WMAR-TV and WJZ-TV.
The WMAR-TV Collection was donated in 1983 to UBalt, consists of 4,000+ reels of 16mm, ~250 reels of 2” Quad, VHS and Umatics; it was the first TV station in Baltimore broadcasting in October of 1947. The WMAR-TV Collection at UBalt spans the time period of approximately 1948-the late 1980s.
Google searchable description of reels in HTML tables which are no where near perfect: these are transcribed from old logbooks from the station: not everything is described; lots of “unknowns” or errors like “Pikesvi11e”; no standardized vocabulary. The best way to search is through specific places, events, dates and people: but make sure you try multiple types of spellings and knicknames!
The WJZ-TV Collection consists of 20,000+ AV items (mostly Umatic and VHS) and was donated to UBalt in 2007. The date of coverage spans from the 1960s-2000. Discoverability is different than WMAR-TV as it is organized by series with PDF inventories whose tape labels are keyword searchable.
Here is an example of one of WJZ-TV’s biggest series, “Eyewitness News”. The only information that we get from the labels of the tapes is the tape and time of the broadcast that is on the tape (while limited, I do still feel extremely lucky that I have at least this).
We also have series-level Finding Aids in CONTENTdm; PDF word searchable but not as detailed as the inventories on the library website and no EAD.
Once you have found an item that you want to provide access to after you have searched website/inventories, then you must ask yourself: What format do you have? Then you can follow the flowchart above to the end result of access in the Internet Archive/ArchivesSpace.
Let’s say that you have one of these [circled in red] formats that needs access. These formats can be digitized in-house!
Betacam, SVHS, Umatic goes to the our video rack, Buffy: The Video Slayer! Which gets captured to our 18TB RAID, configured at RAID 5: which basically means you have quick reliable access to your data (backed up across multiple discs in case of failure the data ban be recovered). Baking needs to be outsourced. I can do basic repair/surgery on cassettes.
Currently access copies only: $30 charge to patron for work (this pays more for my time-reformatting, troubleshooting, QCing the end files-- and maintaining equipment and supplies to clean each video deck)
Compact Cassettes + regular LPs can also be reformatted in-house.
So what happens when you have 16mm film? It gets sent to our Film Station.
It is here at the Film Station where I can inspect, repair, clean in-house (which can save me up to $75/hour).
Then the 16mm gets sent to a trusted and vetted vendor for scanning. The vendor puts all our files on an external hard drive and sends that to us via standard shipping. 2" Quad/1" open reel video/1/4" open reel audio/any other random AV format, as well as any item that needs to be baked, is sent to the vendor as well. For all of these formats to be digitized, the patron pays for all the fees + $30 reformatting fee on UB's end (this pays for my time on the cleaning/inspecting/repairing of the film, maintaining the equipment needed to do this, and then the supplies needed to do this. Quality Control of files from vendor)
From the external hard drives the files are then transferred to our previously mentioned RAID, utilizing Grsync so that no data is corrupted in the transfer.
This is the RAID in the Flow Chart!
From the RAID, the files get uploaded to the Internet Archive using their S3-like API, IAS3 Bulk Uploader: to upload batches of content alongside per-item metadata in an automated fashion. Once it is uploaded and backed up to IA, data is deleted from RAID: backed up, working storage.
We link the Internet Archive item record to an ArchivesSpace item record.
Here is an example of an item-record in AS that links to IA. We are planning to eventually phase out inventories on website and CONTENTdm and use ArchivesSpace only for discovery.
Please contact me with any questions! Thank you.
Managing Production and Access of Digital Collections

Robin C. Pike, University of Maryland Libraries
rpike@umd.edu
SAA, Washington, DC, August 16, 2014
History of Digital Projects

- Early 2000s
- 2005: Office of Digital Collections and Research
- 2009: Digital Collections

The UMD Libraries have been participating in digitization since the early 2000s in Special Collections to fulfill reference requests, tracked with home-grown metadata in an Access database. It later evolved to the Office of Digital Collections and Research in 2005, the first organized effort. This led to the design and implementation of UMD Digital Collections in a Fedora-based repository, best practice standards, boutique digitization projects, and the creation of several collections and projects that would be added to over many years, such as University AlbUM—images and film from university collections. Digitization operations were then moved back to Special Collections in the Digital Collections unit during hard financial times and staff cuts.
History of A/V Digitization

- Films@UM (900)
- Football Films (775)
- The Jim Henson Works (70)
- Ad hoc
- Little mass digitization

Most of the requests and boutique projects from 2000 to 2012 focused on digitizing what was easily achieved—images, publications, and manuscripts—though we did pursue some media digitization. 775 football films were digitized through a vendor from money raised by University Archivist Anne Turkos. The Library Media Services Department pursued licensing agreements with educational production companies and we made almost 900 educational films available to students. The Jim Henson estate donated digital copies of close to 70 of his early works to Special Collections in Performing Arts. These early projects were not created as templates for future moving image digital collections; metadata fields varied between all three collections. Other small projects existed in other areas of the Libraries but did not have consistent practices, staffing, and did not contribute the results of what was digitized to the Digital Collections repository.
In February 2012, the Digital Conversion and Media Reformatting Department (or DCMR) was created in the Digital Systems and Stewardship Division, not within collection areas, to establish a programmatic approach to digitize collections in the seven campus Libraries and Priddy Library at Shady Grove. DCMR seeks to support the Libraries’ collection development goals and strategic priorities for preservation and access by working with collection managers and subject specialists to digitize collections of all formats through a centralized, production-based environment. This has been achievable through administrative-level programmatic and financial support.
One of the first important steps was contributing initiatives to the Libraries’ Strategic Plan. Without these established goals, the initiatives would not have been supported.

After establishing a baseline of existing policies, procedures, best practices, and previous in-house and vendor-based pursuits, I met with more than 20 digitization stakeholders—the people who had been participating in these processes—and then developed an immediate and long-term plan for digitization in the Libraries, focusing on a staggered implementation of growing and new services.

The Libraries had been systematically digitizing image and text materials but based on the stakeholder interviews, we needed to do the same for audio and video formats. In order to implement this ambitious plan, I would need to add equipment and staff, and implement a space and facility re-design.
One of the most important decisions I had to make was what our in-house digitization capacity would be. After consulting with our media collections librarians, and analyzing what legacy playback equipment was available, operable, fixable, and maintainable for a fair price, I ordered the beginning of the hardware and software for three audio digitization stations and one video digitization station. These stations would include specified formats, based on a business case analysis for best use of our available resources, which would fulfill our predicted digitization needs over the next 5 years.

My official policy is that we will either maintain the equipment to digitize the format, or maintain relations with a vendor who can digitize this format within an appropriate period for a patron. We will digitize formats when we can keep the cost down with student labor. We will digitize and deliver these requests within 15 business days, or work with a vendor to expedite formats that we cannot handle in-house. We will not digitize in-house where the expertise is beyond our means (stylus selection for lacquer discs), equipment that is too difficult or too expensive to maintain (maintaining 2” Quad machines costs approximately $20,000 every 500 hours of playback), and when the condition of the material is too complex (mold). We can currently digitize the following audiovisual formats in-house (don’t read list).
My department currently consists of another librarian, who manages daily operations in our on-campus digitization center, and 100 weekly student hours for in-house digitization and quality assurance. Approximately 20 hours/week is spent on in-house audio digitization. We previously had another two-year contract position, who set up the audio and video digitization workstations and developed procedures for these activities.

I rely on the collection managers to perform material selection, which they base on collection usage, predicting patrons’ future needs. While we can bake sticky-shed tapes in our Digitization Center, staff in Preservation and Conservation clean and treat surface mold and other particulates.

I heavily rely on our Metadata Services Department and collection managers to assist with metadata harvest from catalog records, finding aids, or inventories to our adopted metadata spreadsheet for batch audiovisual ingest into our repository. For example, in order to send out our contribution to the American Archive Project, I worked with Joanne Archer, one of our Special Collections Librarians to develop procedures to rapidly create basic metadata records. In three days, three students created 1,600 basic metadata records, enough to send the reels to the vendor with the rest of the selected materials for the project, totaling 7,000 reels. Though stressful, we’ve joked about modifying this guerilla metadata process for future targeted projects. By dividing the digitization duties, we are able to focus more on specific skill sets and are working on larger projects.
Because my department is new, we collect statistics on nearly everything we do so I can justify requests for new resources, like additional student hours, with hard data. For example, using the statistics on the amount of created files and metadata records, and dividing these by the hours worked by my staff in a month, we get the average rate of digitization. Dividing this number by the average of their salaries results in the cost of digitization. I used this rate to predict how long it will take us to complete in-house projects, based on current and proposed staffing levels. I am also using the cost of digitization to calculate whether it will be cheaper to digitize a project in-house or through a vendor.

I've written a slightly longer explanation of this on our division’s blog, which is available here.

I also plan how much digital storage we need for these projects, and incorporate this cost into our annual capital expenditure budget. Since 2012, we've added over 700 audio and video digital assets from in-house and vendor projects. Since reaching in-house production level for in-house audio production last summer, of those 700 assets, we added over 300 digital assets to our repository. Our audio quality assurance and ingest rate is still improving but we have almost 950 in-house and vendor-produced audio assets currently in our QA and ingest queue that we plan to ingest by the end of the year.

We’re also a partner in The American Archive project, a collaboration with the Corporation for Public Broadcasting, the Library of Congress, and WGBH, which will add approximately 7,000 more recordings, or about 9.2 TB some time this fall. While we may be dealing with a lot more TB, we’re dealing with the same scalability issue of ingest, storage, and digital preservation as everybody else.

So far, I’ve focused primarily on digitizing audio but this year, we’re sending 45 films and 100 VHS tapes to vendors for pilot digitization projects. Though this is a comparatively small volume of moving image materials, the massive digital storage required for these projects will inform the future digital storage needs as we progress into moving image digitization, and will help me plan costs for future
After setting up the plan for digitization in the Libraries, I am now focusing my efforts on expanding programmatic vendor-based digitization. I have decided that this approach is the best solution for accomplishing the most amount of digitization over the next 10-15 years while the materials are still viable.

I plan approximately 10-20 in-house digitization projects in one year, staggering production starts. I am also managing 13 vendor-based mass-digitization projects during FY15. I have been successful to expand digitization projects from the Special Collections and University Archives and Special Collections in Performing Arts departments into the Architecture and Art Libraries, Humanities and Social Sciences collections, Government Documents, and Engineering and Physical Sciences Library collections. Not all of these projects contain audiovisual materials, but we will be focusing on four outsourced and at least one large in-house audiovisual project over the next year. I continue to meet with a growing list of digitization stakeholders in the Libraries (37 for FY15 projects) to make sure they are aware of our expanding services, to collect projects for the coming year, and to help them plan and implement these projects; outreach to staff about programs and services is just as important as outreach to patrons.
In addition to planning and pursuing FY15 projects, I started to plan FY16 projects this past June. Planning two years of projects at one time is a bit crazy and it’s driving my stakeholders crazy, too. However, we need to be planning this far in advance so we can allocate or plan to pursue enough funding to support the projects as part of the project planning cycle. And I’m not just planning to pursue funding for one or two projects; we need to fund every prepared, prioritized project, and pursue a lot more funding for audiovisual digitization projects.
To try and keep everything straight, and avoid spreadsheets, I’ve started to use software like Visio and OmniGraffle to plan workflows, and Microsoft Project to plan production timelines. I don’t have everything planned yet, but I am working towards creating a two-year workplan for all projects, both vendor and in-house.

A key aspect of managing programmatic, sustainable digitization projects, especially complex audiovisual digitization projects, has been creating new workflows. In documenting these workflows, I’ve been able to work out some inefficiencies (like making we have all the information we need before we get a vendor estimate, such as including footage if selecting film reels). These particular workflows represent the main processes for all digitization projects, though it is specifically adapted for vendor-based projects where external funding is required.
And this is where some people may start to think, “I don’t have any money for outsourced digitization.” Prior to FY15, any audiovisual materials we have sent out to be digitized have been paid either through the Preservation and Conservation Department’s budget, or someone’s gift account if they had access to an account, so essentially, I asked other people if I could use “their” money. This fiscal year, we tried something a little different—a committee!

The Digitization Initiatives Committee is comprised of five members, one member from each Library division, including our Development Officer. To fund FY15 projects, we solicited digitization project proposals, which later became part of the project plan, created estimates, and located funds within several collection-area operating budgets, and gifts and endowments accounts. Of the 13 projects, only 4 are audiovisual digitization projects, but they do comprise over $31,000 of the overall budget.

This method of funding is not sustainable because some of the accounts are non-renewable. For FY16, we’ve moved the proposal process to this fall because we’re going to get full vendor estimates before we try to find funding so the estimates are closer to real costs, and so our operations budget requests fall within the administration’s operation budget planning schedule; by default, we may get something. I don’t take “there’s not enough funding” or “other divisions don’t want to give up their operations budget” as a “no.” operations budget” as “no.”
To expand our buying power, digitizing more material for the amount of money we have, I have taken advantage of our Lyrasis membership. Lyrasis has negotiated slightly lower digitization rates for their members with a variety of digitization vendors and possesses better buying power with many members than UMD does alone. All of our mass-digitization audiovisual projects will be digitized by Lyrasis contracted vendors this year. Arranging mass-digitization this way also allowed me to write one sole-source agreement instead of sending out multiple requests for proposals per format. Despite selecting these vendors, I continue to set meetings at UMD and at local vendor locations to talk to other vendors, learn about their services, develop these relationships, and look for better rates or added-value services, like embedding metadata.

I’ve created a set of standard technical specifications for audio and moving image deliverables, which is made easier by using the same vendor and building that relationship. I’ve found that when we’ve built a relationship, we have a much lower incidence of errors and in-house quality assurance goes much more quickly and we have to ask for less re-work. If I’m not sure about a standard or process, I ask the vendor what other cultural heritage institutions are requesting, for example, “What are leading institutions asking for in terms of embedded video metadata?” or “How do you recommend packing and shipping 45 reels of 16mm film?”
As we expand beyond four or five major audiovisual digitization projects in a year, listening to every recording for the entire duration will be impossible. In order to identify audiovisual media to send it for digitization, surprisingly basic metadata is needed—a title, date, unique identifier, and some format information will usually suffice—but, to make these recordings useful, they need to be searchable. One of the biggest issues with digitizing audiovisual media is that most media has minimal metadata, and making recordings fully accessible requires actively-listening staff creating descriptive metadata of content and timestamps. Over the next year, my colleagues and I will experiment with several automated audiovisual indexing and transcription tools and services, such as Hipstas, Pop Up Archive, or ResCarta, to name a few, with the hope that in the future, we will spend less time generating metadata to make our materials searchable.

I also want to pursue other funding resources with our collection managers so we can use our funding to match that of granting agencies, and potentially double our production.

The Libraries have pursued joint digitization projects with other institutions and organizations for our French Pamphlet collection and government documents, and are considering doing the same for some of our larger performing arts audiovisual collections.

You can find out more about current DCMR projects on the Digital Systems and Stewardship division blog.
Contact Information

Robin C. Pike
Manager, Digital Conversion and Media Reformatting
rpike@umd.edu
@AVArchivist

Thank you!
Preservation Services department
→ in-house conservation and preservation reformatting of rare & unique AV
→ preservation consultation for new born-digital production production projects
  → Specifically content for which the Library will be a logical access point
>>> DEPT HISTORY
• UVa Lib Preservation department only formally came into being in 2005
  → lots of work in last several years to centralize preservation services, AV focus part of this trend
• Library began to make preservation a development focus
• Preservation dept. Mellon Foundation grant 2008
  --> funded Book & Paper Conservator, fitted out conservation lab out, as well audiovisual equipment
~2010 AVPRES STATE OF PLAY

- Preservation Services owned almost no audiovisual equipment
- No specialized staff to oversee reformatting by vendors
- Little useful metadata for rare & unique audiovisual holdings

- Any digitization was dependent on staff labor in the Library’s Digital Media Lab (student & faculty AV production space)
  - transfers were not archival quality
  - only other option was researcher paying for outside vendor

- OUTSOURCED REFORMATTING?
  - didn’t have strong technical guidance for target formats

- VAGUE METADATA for archival AV
  - collection-level MARC records w. difficult to decipher descriptions of AV holdings
  - Spec Coll sometimes “shadowed” records for AV holdings within MSS collections
• Where to start with AV preservation practices?
  → where are WORKFLOW GAPS preventing access?
  → where can effort be leveraged, biggest impact be made?

• INFRASTRUCTURE for digitization to meet RESEARCHER REQUESTS
  → reasonable starting point, with immediate positive impact

• BUILD OUT of format-specific AV preservation workspaces and in-house reformatting labs
• Workspace for safe handling and inspection of film
  –can perform basic conservation actions: ASSESS CONDITION and REPAIR
  –INFRASTRUCTURE needed to create item-level condition reports, PREP FILM for OUTSOURCED lab reformatting
• DIGITIZATION lab for most commonly held audio formats
  → grooved disc formats (LPs, also more arcane formats like transcription discs), audiocassettes, and ¼” open reel audio
• DIGITIZATION of video formats
→ ¾” Umatic (and Umatic SP), VHS/S-VHS/VHS-C, Betacam/BetaSP/Digibeta
• you CAN’T digitize all formats IN-HOUSE!
  ➔ prioritize collecting equipment that represents what you have in your collections!
  ➔ even if you plan to mostly outsource, PLAYBACK EQPT. often essential to IDENTIFY
    WHAT CONTENT you have

• KNOWLEDGE: what AV do we have in our collections?
  ➔ creating granular item-level inventories of legacy AV collections...
  ➔ …capturing preservation-minded fields! PRECISE FORMAT INFO (track config for
    ¼”), runtime/stock length

  ➔ can loosely prioritize by AGE of format! Some formats are just bonkers and
    trending towards obsolescence much faster
    ➔ playback equipment will soon be unavailable for the oldest/weirdest
    formats (Mike Casey of IUB says ~12 years)
“Many analog audio recordings must be digitized within the next 15 to 20 years — before sound carrier degradation and the challenges of acquiring and maintaining playback equipment make the success of these efforts too expensive or unattainable.”

– Library of Congress National Recording Preservation Plan

• DIGITIZATION, not conservation actions
  ➔ it’s not worth REHOUSING magnetic media
  ➔ we’re well past the point where preventative conservation actions are reasonable plans of action for magnetic media

• unlike possibly “obsolete” digital file formats, playback technology will likely never be reverse engineered/hacked
  ➔ emulation is not a solution for most audiovisual formats, which makes them a more urgent priority
NEED EQUIPMENT?
• OBsolescence alarms ringing, now is the time to hoard audiovisual playback equipment
  → eBay! Reasonably priced, and semi-functional decks can be refurbished

• BEFRIEND any audio/video engineers at your institution!
  → years of production experience = familiarity w. what we now term legacy formats
  → technical expertise to maintain AV playback equipment is precious!

• SCAVENGE LOCALLY!
  → locate and rummage through the AV equipment graveyards at your place of work!
We’ve built in-house reformatting facilities, but they don’t yet effectively SCALE
→ our current reformatting labs are for 1-to-1 transfers, no mass digitization
→ I am only staff member working on legacy AV reformatting

>> Trying to address how we could train other staff or student workers to do transfer work
→ important to articulate to administrators what technical work is appropriate for skill/education levels
→ building new, lighter-weight digitization stations that could travel
→ only for formats that require less-specialized handling skills
→ AV formats for which reformat via DATA TRANSFER as opposed to SIGNAL CAPTURE: DVCAM, mini-DV
   → soon to pursue DAT reformatting workstation using DDS tape drive

>>> LABOR ISSUE
• Student workers? (Repeatedly suggested re: transfers)
   ---> as a specialist, up to you to make case for specialized professional staff.
   ---> ETHICAL ISSUE: for both audiovisual archivists and the integrity of your content
• briefly touch on BORN-DIGITAL audiovisual content
  → Biggest point of departure from traditional analog workflows is concept of STORAGE SPACE (“Where do we put stuff?”)
  → most contemporary AV production workflows don’t write to a tape-based medium

• Music Dept Concert recordings…moving them away from creating optical discs.
• A-School visiting lectures that used to circ on VHS/DVD
>>> LET’S PLEASE STOP CREATING PHYSICALLY-EMBODIED MEDIA
   → Getting over the reassuring aspects of a “tapes on the shelf” model of stewardship, MOVE TOWARDS DIGITAL CURATION
   → creating new optical media/reformatting to videotape is creating tomorrow’s backlog
   → Physical media are “dead” on shelf, need to be handled, identified. They require signal extraction technologies to get their content off the carrier.

   → ACTIVELY MANAGED DIGITAL OBJECTs afford many conveniences over objects on shelves...IFFF we have TOOLS!
      → Obviously this is where need for IT support is needed, and where ANXIETY is FELT
      → Production workflows that might have been SELF-SUFFICIENT w. physical media likely NOT when born-digi
      → If we support production w. proper tools, NO ONE SHOULD MISS optical media/tape-based workflows!
</rant>
~2012

>>> EXTEND TOOLS AND SERVICES to born-digital workflows
  → Focus on bringing existing “in-house” production workflows into a managed preservation environment

  → Existing repository ingest workflows for text & image only,
  → Points of repo access/tools limited to certain departments
  → Local production workflows that were self-sufficient for analog, NOT as born-digi!
• Avalon! Hydra application for managing audio and video
  → developed by IUB/NWU under IMLS grant
  → pursuing some different local implementations: manage MASTERS in Fedora 4, metadata customization to use RDF-based standard

• Moving towards goal of democratic access to a preservation repository
  → Avalon will let define/restrict user roles so that we can open up system to local content managers

• Looking to collaborate w. Hydra partners to create pipeline to push descriptive metadata into ArchivesSpace
THNX!
steev@virginia.edu
Contact Information

- Siobhan Hagan
  - shagan@ubalt.edu
- Robin C. Pike
  - rpike@umd.edu
- Steven Villereal
  - steev@virginia.edu

Thank you!