

**UC Audiovisual Preservation Task Force
Report and Project Proposal
August 20, 2014**

Task Force Members

Barclay Ogden, Lead (UCB)
Stephen Abrams (CDL)
Dawn Aveline (UCLA)
Roger Smith (UCSD)
Patricia Smith-Hunt (UCR)

In May 2014, the UC Strategic Action Group 3 (Collection Building & Management) (SAG3) issued a charge to convene an Audiovisual Preservation Task Force to examine the issues surrounding a system-wide approach to preservation and access of our audiovisual collections. The charge of the AV Task Force includes:

1. Conduct an **assessment** of the UC Libraries' holdings (e.g. locally recorded media that document the intellectual life of our campuses, versus commercially produced media which represent significant historical investments) taking into account the variance among campuses.
2. Briefly articulate the **extent of the problem** and the real impact if no action is taken. Propose a **prioritization strategy**.
3. **Analyze and recommend possible solutions** for providing preservation and access to recordings, weighing in on formats, metadata access and the infrastructure requirements to sustain a UC media preservation program (e.g. coordinated through a single campus' leadership, multi-campus collaboration, or CDL-centralized approach).
4. Develop an intellectual property rights **framework** or assessment checklist to aid campuses in assessing which parts of their media collections can be digitally reformatted to enable both preservation and access.

Deliverable

The Audiovisual Preservation Task Force will submit a report documenting the scope of work recommended and expected outcomes of a preservation project to SAG3 by August 29, 2014. Based on SAG3's review, this membership may evolve to a Project Team to carry out the project outlined.

Extent of the Problem

Film, video and audio materials have enormous potential instructional and experiential values, making the past feel present, alive and tangible. With media permeating modern life and media literacy rising—in 2013, according to the Cisco Visual Networking Index, video represented 66% of global traffic on the internet and is anticipated to jump to 79% by 2018—users expect to find audiovisual materials along with textual and photographic materials.

There is a real danger that classroom and distance learning/further education experiences – as well as scholarly research and publication processes and cultural heritage experiences – will lose relevance for students, educators, and the public if they remain walled off from the wider world of moving images and recorded sound.

— Paul Gerhardt and Peter B. Kaufman, *Film and Sound in Higher and Further Education* (The Higher Education Funding Council for England, 2011).

UC Libraries have large holdings of moving image and sound recordings: as of the 2011/12 UC Library Statistics, a total of 1,121,863 recordings in analog and digital formats with an estimated asset value (for insurance purposes) of \$391 million. Roughly 800,000 or 70% of these recordings are in analog formats that require playback equipment less and less available in UC libraries. Moreover, time is running out for conversion of analog media; some estimates predict that in less than two decades most analog magnetic media will have deteriorated beyond playability and, consequently, ability to be digitized. As a consequence of limited functioning playback equipment for analog recordings and deteriorating media, UC needs to take action now to ensure continuing access to audiovisual research resources with permanent research value.

The Task Force Process

The Task Force conducted its business via several conference calls and email communication to identify issues to be addressed, actions to be taken, and to scope out phases for a pilot project. Task Force members contacted the UC Collaborative Collection Development via Digitization Task Force to leverage its insight about collection priorities and funding approaches, two ULs to open investigations into funding, and CDL staff to scope out technical issues and costs regarding access to and preservation of digital files of audiovisual recordings.

The Project

The Task Force proposes a pilot project to identify UC's analog audiovisual holdings with permanent research value, assess their preservation and access needs, establish an intellectual property rights framework with which UC can determine which recordings can be digitally reformatted for preservation and online access, and develop a production strategy to achieve that goal.

The pilot forges a path of scalable phases that adapts to changes in levels of funding and staffing. Options for project staffing include adding UC staff or contracting out project work. Options for funding in addition to UC libraries' funds include federal, state, and foundation grants.

The pilot includes:

- selection criteria for prioritizing recordings to which to provide future access
- framework and best practice for due diligence on intellectual property rights
- 3,000 high priority analog recordings converted to digital formats
- guidelines for metadata for audiovisual recordings
- minimum required metadata for all recordings converted
- online access to all recordings for all UC campuses

- preservation of the digital files and associated metadata
- \$1,000,000 pilot cost, estimated as follows:
 - 3,000 recordings @ \$200/recording
 - CDL infrastructure development*, \$90,000
 - 2 FTE project staff @ \$75,000/FTE/yr.
 - Processing and inspection equipment, \$10,000
 - Campus contributions of time for assessing needs, selecting titles, performing quality control for digital recordings, and creating/upgrading metadata as needed.
 - 24 months following receipt of funding

Pilot Project Objectives

1. Secure funding for pilot project. (Investigate campus libraries, UCOP, and grants. If full funding cannot be secured, a scaled back pilot project to preserve fewer recordings could be undertaken in the same two-year project period.)
2. Identify and adopt selection criteria for titles to be preserved for continued access. (Engage collection development staff to determine selection criteria; possibilities include uniqueness, rarity, demand, value, risk of loss, and intellectual property limitations.)
3. Using selection criteria, determine the scale of UC need, cost out the components for an analog-to-digital conversion project at scale, and identify cost effective strategies for ongoing services.
4. Identify and cost out the components for an online and/or off-line infrastructure to manage, serve, and preserve the digital audiovisual files at scale.
5. Identify minimum metadata needs and costs.
6. Address potential intellectual property issues, determine limits to access, and develop best practices for due diligence.
7. Determine requirements for digitization and quality control of the analog-to-digital conversion process (as well as requirements and costs for closed captioning).
8. Digitize 3,000 recordings of high value to UC for preservation and online access for all campuses, assuming a \$1 million project budget.
9. Identify campus services to provide ongoing digitization as needed following the full-scale project (likely based on campus digitization services participating in pilot project).
10. Identify sources of funding for conversion, hosting, and ongoing life cycle management of recordings at scale.
11. Evaluate potential and cost effectiveness of the pilot to be scaled to meet the needs of UC libraries.

Pilot Project Work Plan

The activities needed to successfully complete the project draw on the expertise of all campus representatives. The following phases are used to process, preserve, and provide online access to 3,000 recordings from the ten campuses over a two-year pilot project.

Phase 0: Planning. Months 1-6.

Pilot project funding is secured. The project production path is planned, including assessment, selection, digitization, quality control, preservation storage (at least on a pilot basis), and uploading to access servers. Metadata, IP, digitization, and preservation issues are addressed.

- Secure pilot project funding. (project team)
- Identify pilot project staff and campuses that wish to do some of the conversion work. (project team)
- Develop holdings assessment tool. (project staff and campuses)
- Develop selection criteria. (project staff and campuses)
- Develop and disseminate RFQ for conversion work. (project staff)
- Develop and deliver workshops/webinars and tutorials to UC libraries to cover selection, cataloging, and quality control. (project staff)
- Identify digital infrastructure requirements. (project staff)

Phase 1: Assessment. Months 4-6.

In coordination with campus representatives, determine scope and scale of needs of audiovisual collections with permanent research value. (project staff)

Phase 2: Selection & Preparation for Digitization. Months 7-15.

- Develop infrastructure to manage and serve audiovisual access files. (CDL or other digital repository service)
- Select recordings to be digitized and implement due diligence procedures regarding rights. (campuses)
- Award digitization contracts to vendors. (project staff)
- Prepare the recordings for digitization
 - Provide preliminary descriptive metadata and rights information for each recording. (campuses)
 - Ship original recordings to the project location and upload metadata to CDL. (campuses)
 - Review supplied metadata and add metadata as needed, including technical notes about the original materials. (project staff)
- Ship analog originals to the digitization service provider for inspection, preparation, and digital transfer. (project staff)
- Write interim report for SAG-3 after one year. (project staff and project team)

Phase 3: Digitization/Digital Preservation. Months 13-24

Several sub-phases happen almost concurrently. The materials are reformatted to digital; the digital files are QC'd and packaged for preservation; access files are uploaded to servers.

- Digitize recordings and return originals to campuses. (vendor)
- QC digital files. (project staff and campuses)
 - Undertake technical quality control. (project staff)
 - View recordings to report problems with integrity or readability, and improve discovery metadata after having viewed the recordings. (campuses)
- Upload access files and associated metadata to access server. (project staff)
- Upload preservation files and metadata to digital preservation repository. (project staff)

Phase 4: Access & Pilot Project Evaluation. Months 19-24.

What worked? How much did it cost? How is the online audiovisual resource used? What are the resources and skillsets needed to progress from pilot to full production? How is the service continued after the full-scale project is completed?

- Survey campuses on the time and resources required to participate in the project. (project staff)
- Survey users on values and uses of online audiovisual recordings. (campuses)
- Evaluate the pilot for expansion to full-scale. (project staff and project team)
- Investigate funding sources to support full-scale project. (project team)
- Recommend options for continuing conversion services following a full-scale project. (project staff and project team)
- Write final report on the Pilot Project for SAG-3. (project staff and project team)

Recommended Next Steps

1. Confirm administrative interest in proceeding with the pilot project.
2. Establish a pilot project team to undertake the following:
 - a. Investigate opportunities for funding from UC libraries.
 - b. Investigate grants from foundations and government agencies.
 - c. Investigate options for contract services, consultants, and for UC staffing.
 - d. Assuming success with funding, implement pilot project.
3. Charge the members of the Task Force to become the pilot project team following discharge of the Task Force.

*** CDL Infrastructure Development**

The recently announced UCLDC digital asset management system (DAMS) is intended for management of objects (including audiovisual assets), i.e., organization, metadata creation, derivative creation, etc. Notably, the DAMS is not a preservation repository; DAMS users will have the option of selecting objects to send to Merritt for long-term preservation.

The UC Curation Center (UC3) at the CDL operates the Merritt repository as a shared resource available to the UC community. Merritt already is being used to manage over 34,000 audiovisual resources totaling approximately 4.6 TB. All digital assets managed in Merritt are geographically replicated (currently three replicas, soon to be expanded to five) and subject to ongoing fixity audits to detect and correct local bit-level damage. Costs for preservation storage and management are based on the number of TB of data stored on its servers.

At present, the options for delivery of audiovisual assets are limited to HTTP file-level retrieval and retrieval of multi-file assets encapsulated in a single Tar, Gzip, or Zip container; no streaming delivery currently is available. Support for streaming delivery could be investigated as an addition to Merritt's service functions with implementation subject to UC3 and CDL's other ongoing and planned activities and initiatives. The highly distributed nature of Merritt's architecture, with processing centralized at the UCOP Oakland data center but preservation storage hosted at UCLA and UCSD/SDSC, could introduce significant, and potentially unacceptable latency into media streaming. Alternative approaches could be investigated, such as co-locating access files and streaming servers at a commercial cloud provider.