



NDSR Project: Harnessing Media Micro-Services for Stewardship of Digital Assets at CUNY TV

Goal Summary

This project is designed to assess the developing state of archival procedures of the CUNY TV Library, including those that pertain to media acquisition, storage, digitization, information management, and content dissemination. The resident's assessment shall lead to reporting which proposes a prioritized list of advancements or adjustments to be pursued or implemented during or after the residency. The resident shall then be supported in enacting selected recommendations through re-design or new development of open source software and documentation. Some of the implementation work of the resident is pre-defined in order for the library to prepare to meet upcoming preservation challenges and further the integration of information management systems.

The project would work specifically with the micro-services currently in service in CUNY TV's Library. These micro-services are primarily comprised of bash scripts and defined workflows to accomplish specific tasks, such as transcoding, assessment, delivery, storage, metadata harvesting, logging, and digitization. The project anticipates coding, technical experimentation, feature development, and work on the library's databases in order to better align these workflows with preservation standards, local needs, and digital sustainability. Because of the technical nature of the objectives both the project mentors and CUNY Television technology and engineering staff are available to work closely with the resident in support of these objectives. Additionally the project mentors gained valuable experience in conducting a version of this project on a smaller scale in 2013 with Athena Christa Holbrook, an intern from the NYU Moving Image Archiving and Preservation program.

To create more efficient and well defined workflows for preserving, maintaining and providing access to CUNY TV digital assets, the resident would also evaluate the data integrity of our assets in preparation for an anticipated data migration from LTO-5 to LTO-7 tape in 2016. This migration will give us the opportunity to update our LTO and packaging workflows by interviewing employees in the library, broadcast and the IT department to find out how the use of LTO and overall approach to digital stewardship may be improved.

Specific Objectives

To interview pertinent staff members about LTO and media micro-services and make suggestions to the archivist on how to improve the system.

- To develop a work plan and actionable micro-service supporting the migration of digital assets from LTO-5 to LTO-7.
- Analyze the data integrity of our digitized collections along with our born digital assets including analyzing checksums and defining the handling of duplicate files.
- Devise a work plan to implement CUNY TV's digital assets into our new Digital Asset Management system, ResourceSpace.
- Review and document existing micro-services, including scripts that facilitate content acquisition, digitization, packaging, transcoding, delivery, and quality control.
- Interview the archive's operators to determine and prioritize needs and opportunities for development, with a focus on the implementation of externally defined archival standards.
- Design and implement procedures or scripts to gather process history metadata and acquisition context during acquisition procedures.

*Timeframe &
Deliverables*

Month 1

- o An introduction to the mission, work, and collections of CUNY Television.
- o An introduction to the CUNY Television library, its staff, services, and responsibilities.
- o Introduction to relevant departments including the New Media, Broadcast, Programming/Scheduling, and IT.
- o Shadow the broadcast librarian and train on content acquisition, registration, and dissemination.
- o Interview select staff about library operations including the Control Room Supervisor, Broadcast Librarian, Captioning Manager, and Archival Technicians.
- o Under the guidance of the broadcast librarian and control room supervisor, review operational procedures between the library and broadcast groups, file delivery, transcoding, quality control, and record-keeping.
- o With the project mentors, get an overview of the library's micro-services and datasets, along with their integration and use.

Month 2-3

- o Interview representative members of New Media, Broadcast, Programming/Scheduling, and IT departments regarding accessibility, library services, and information and media transactions with the library,
- o Migrate current micro-service scripts into an organizational GitHub account. Update and document installation procedures (a Homebrew tap) and dependencies.
- o Use a GitHub issue tracker to detail bugs or enhancement requests pertaining to the micro-service collection.
- o Utilize an internal trac system implementation to centralize and coordinate internal documentation of archival procedure and assignment of tasks to staff.
- o Draft a strategy for managing process history metadata.
- o Implement process history metadata into the library's digitization micro-services (vrecord, paperingest, and others).

- o Assess the use of each micro-service, the state of its code and documentation, and feature needs.
- o Review and update internal documentation of each micro-service.

Month 4-5

- o Document CUNY TV's existing archival information package (AIP) implementation and propose refinements.
- o Enact selected AIP refinements.
- o Coordinate development of a micro-service to assess compliance of an AIP to the created documentation.
- o Standardize transcoding micro-services and integrate process history strategies into relevant micro-services.
- o Document a workflow for migration of the existing LTO collection (approximately 300 LTO-5 tapes with a redundant set) to a potential use of LTO-7 in 2016.
- o Address feature requests and bug fixes for micro-services.
- o Provide an interim progress report on the progress of the residency.

Month 6-7

- o Design an automated testing environment for the micro-service collection in order to enable regression testing.
- o Run a test implementation of the designed LTO migration workflow, covering 10 LTO-5 tapes of various content types, including data verification and packaging assessment procedures.
- o Address feature requests and bug fixes for micro-services.
- o Review and revise process history implementations.
- o Apply AIP compliance micro-service to select collections, document AIP antidote procedures or fixes that may be needed.
- o Update AIP compliance micro-service accordingly.

Month 8-9

- o Draft final progress report culminating the discoveries and progress of the residency and outlining the remaining recommendations CUNY TV after the residency has been completed.
- o Document results of LTO migration tests and provide recommendations for scaling it to the entire data tape collection, including timeline and assessment criteria.
- o Selective implementation of micro-service features.
- o Coordinate and draft public documentation on the micro-service set and project results.

Deliverables

- A report analyzing the data integrity of our digitized collections and born digital assets. A work plan regarding the transition to LTO-7.
- Updated micro-services and associated training.
- Create a user manual for our media micro-services that defines and documents the various uses and best practices for CUNY TV's micro-services scripts.
- Propose an integration plan for LTO records and the respective metadata into our existing FileMaker database.
- Research and develop a workflow that integrates our digital assets into ResourceSpace, our Digital Asset Management system.

*Resources
Required*

CUNY Television will provide the resident with a new Mac computer and workspace based in the library, which provides an open office working environment amongst the project mentors and other key staff. The resident will also have access to a workstation in CUNY TV's data center and control room. We will ensure that the resident has sufficient hardware access for the assignment, including access to relevant networks, systems, and databases as well as decks for digital media formats such as LTO, XDCam, DV, and various digital recording formats. The resident will be supplied with installations of required software such as Final Cut Pro, FileMaker, FFmpeg, MediaInfo, and installation of CUNY TV's micro-service collection with all included dependencies. The resident will also be supplied with a station passkey and business cards.

The project mentors as well as technology staff will also actively support, train, and make resources available to support the resident's work within areas of micro-service design, audiovisual assessment, transcoding, and OAIS implementations.

Context

The library at CUNY Television has long served as a hub of media workflows. Newly produced or acquired content is submitted to the library for registration and cataloging and the library subsequently ensures that the content is accessible for reuse, broadcast, online access, or other uses. Formerly the library required these media transactions in videotape formats, but in 2011 the library began transitions that enabled and encouraged the submission and dissemination of digital file-based media. Also in 2011, CUNY Television expanded the staffing and services of the library to include a preservation program, research services, and new focuses on information management, media accessibility, and preservation standards.

In order to accommodate a large and steady rate of data and processing from both acquisitions and digitization efforts, the library designed an OAIS-inspired packaging standard for content management and adopted LTFS formatted LTO-5 tapes as a digital storage medium. Whereas the library formerly managed circulation of videotapes, the library developed systems to evaluate incoming media and to transcode it to the requirements of particular uses. These operations along with other tasks (including quality control, delivery, and reporting) are performed through the application of micro-services or sets of micro-services on archival packages. Locally these micro-services exist as registered bash scripts that coordinate open source tools such as FFmpeg, MediaInfo, ExifTool, and md5deep.

Presently, CUNY TV's internally developed media management and processing micro-services have been successfully operating for several years during a time of ongoing development and enhancement. This effort has streamlined archive and access workflows and brought an open source approach to broadcast television. At this time, we are pursuing an archival workflow re-design phase where the existing micro-services are evaluated, documented, and refined. This phase would include developing workflows for upcoming challenges such as LTO data migration, finishing and deploying a fixity registry, and extending opportunities for access and reuse of content. We believe the re-design will further our long-term digital preservation goals.

Required Knowledge and Skills for Resident

Applications for the residency should have a graduate degree or equivalent in Library and Information Science, Moving Image Archiving and Preservation, or Archival Studies. Additionally the resident should have the following knowledge or qualities:

- Interest in digital archiving
- Ability to work independently or collaboratively
- Detail oriented and highly organized
- Project management skills
- Familiarity with PBCore
- Familiarity with FFmpeg
- Willingness to learn emerging digital technologies
- Familiarity with the Open Archive Information System (OAIS) framework

Preferred Knowledge or Experience

- Experience working with born-digital materials
- Knowledge of audiovisual file formats
- Experience working with bash scripts and XML
- Experience in the PREMIS metadata standard or handling process metadata
- Knowledge of FileMaker or MySQL