Provenance, Journal of the Society of Georgia Archivists

Volume 34
Number 1 *Audiovisual Issue*

Article 11

10-2016

"Is This Enough?" Digitizing Liz Lerman Dance Exchange Archives Media

Bria Parker
University of Maryland - College Park

Robin C. Pike
University of Maryland

Vincent Novara
University of Maryland - College Park

"Is This Enough?" Digitizing Liz Lerman Dance Exchange Archives Media

Bria Parker, Robin C. Pike, Vincent J. Novara

A modern dance performance leaves an impression on the memory that is hard to place into words after the fact. The unique and powerful use of the human body; the sound design that provides a semblance of time or atmosphere; the narrative or meaning that movement conveys; the costumes, set pieces, and lighting that complete the visual presentation: all of these elements contribute to a performance event that is unlike any other art form—one you must witness to fully experience. This need for attendance, or presence, creates a significant challenge to the dance scholar who researches a dance work years after the performance. Researching a dance company presents different challenges, as companies are both an organization framed by the necessity of day-to-day operations, and a creative collective looking to engage communities in new and exciting ways. The archival dance collections found at Special Collections in Performing Arts (SCPA) in the Michelle Smith Performing Arts Library at the University of Maryland (UMD) support these types of research.

SCPA was established in the early 1970s, when multiple archival music collections and research centers placed in what were then the Fine Arts Library and the Department of Music merged into a singular entity, Special Collections in Music. Early collections included the Music Educators National Conference Historical Center, the American Bandmasters Association Research Center, and the International Clarinet Association Research Center. By the 1990s, the scope had expanded to the performing arts, and the curators pursued collections in theatre and dance.

The dance collections may contain any form of material resulting from the development of new works or from the logistics of performance. On the creative end, these materials include dance notation, rehearsal notes, videography used for reference in the creative process, audio accompaniment on obsolete formats, costume sketches, and contextual information that inspire a work. For the day-to-day operations of a dance company, there are ledgers, rosters, tour itineraries, promotional photographs, professional correspondence, brochures and advertising, grant materials, videos documenting performances, and a myriad of objects that contribute to institutional memory. The Liz Lerman Dance Exchange Archives is an ideal example of a collection of an internationally recognized dance company, but also includes the archival record of the visionary founder.

Overview of Liz Lerman Dance Exchange Archives and the company (02:25)

Liz Lerman Dance Exchange Archives is one of the rare collections at SCPA of an active dance company. Based in Takoma Park, Maryland, Dance Exchange exists to engage and educate audiences of all generations in making art through dance. The company's creative vision came from its founder and namesake, Liz Lerman. She nurtured and directed the company for the first thirty years before stepping down in 2006 to pursue new ventures in dance. During residencies and international tours, the company, under Lerman's direction, was guided by four questions: Who gets to dance? Where is the dance happening? What is it about? Why does it matter? The Dance Exchange Archives is also the largest dance collection in SCPA, presently spanning over 150 linear feet in papers, and containing 1,333 video items and 927 audio items. SCPA acquired the collection in 2004, and the original transfer of papers was processed over the next two years. The audio and video, though described minimally at the item level, were given very little preservation attention apart from storage and monitoring in a stable environment.

Overview of the Liz Lerman Dance Exchange Archives acquisition history and predominant formats (01:01)

From the earliest acquisition-related discussions with the company, the digitization of the video assets was indicated as a priority for both parties. A grant from the National Endowment for the Arts allowed for the hiring of a project archivist to process and describe what was then 130 linear feet of papers in 2005. Concurrently, the Dance Heritage Coalition provided an intern to review and describe the contents of the audio and video items included in the collection. Despite the steps taken, it took over eight years to initiate a digitization pilot due to a lack of funding and the unavailability of technically capable staff to plan the project, and it will take several more years to digitize the entire video collection at current funding levels.

Until UMD Libraries established the Digital Conversion Media Reformatting (DCMR) unit, few digitization protocols were in place and procedures for digitizing video assets were not in development. Despite the existing standards from the profession and model workflows at peer institutions, UMD needed to undertake a pilot in order to identify the best implementation protocol for digitizing and making video accessible. The Dance Exchange video items provided the ideal format and content for such a pilot.

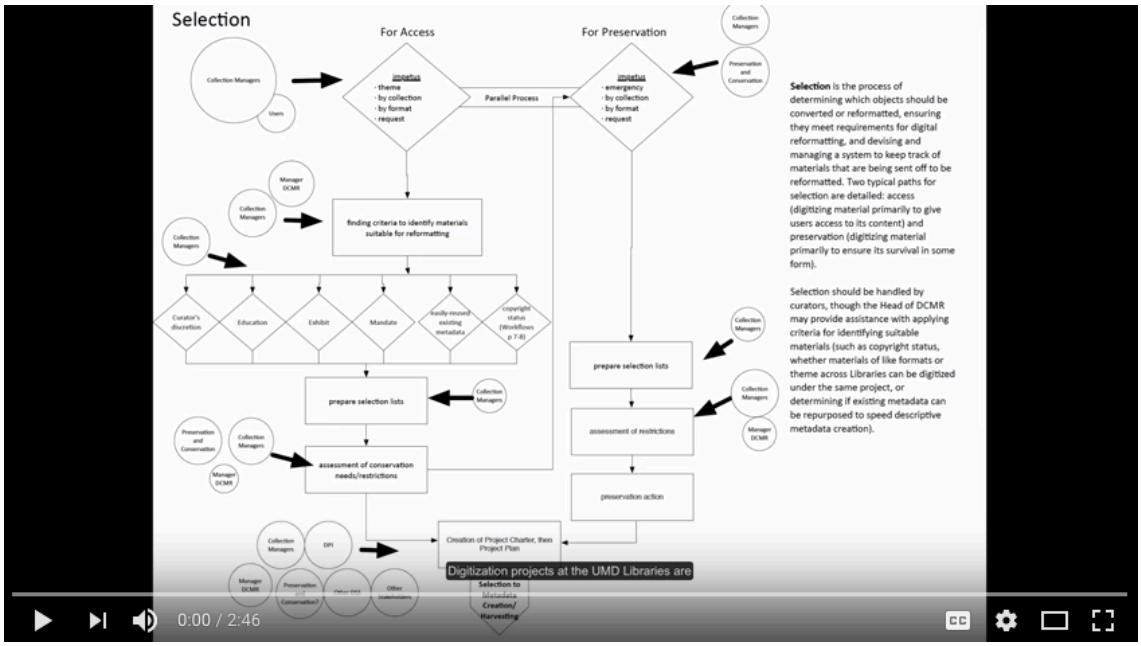
Working with a vendor

University of Maryland Libraries manages digitization across the seven campus libraries through DCMR. Small projects and most patron-based requests are routed through an in-house digitization center, though the department cannot currently digitize video and film due to the complexity of the setup required. Because UMD is a member of Lyrasis, a non-profit community for libraries, archives, and museums, discounts are available through their Digitization Collaborative partner vendors. Throughout 2015, the manager of DCMR worked with these vendors on a series of technical specifications

that are now applied to large-scale projects, such as Dance Exchange, as well as to small requests that are digitized as needed.

Projects at the UMD Libraries are funded through a proposal process in order to focus limited funding for digitization across the Libraries. The Dance Exchange pilot, consisting of 100 VHS cassettes—or 13 percent of the video items—was selected via this process as a part of a larger Libraries-wide initiative to prioritize the digitization of deteriorating audio, video, and film, and to establish workflows for this process. This pilot was the only video digitization project slated for fiscal year 2015, and it helped to establish the workflow for creating minimal description for vendor-based audiovisual digitization projects. Funding was provided by the "Robert Smith Fund for the Performing Arts Library" gift account. Digitization for the Dance Exchange pilot project was contracted to a vendor as UMD did not have the necessary in-house personnel expertise and extensive equipment needed to complete the project. Additionally, the pilot was scheduled to conclude within one fiscal year, faster than a single technician's output while operating one machine.

Working with a vendor involved negotiating a Statement of Work (SOW) that outlined the technical specifications, deliverables, and costs of the project, as well as terms and conditions. The negotiation process took more time than previous SOWs negotiated by the DCMR manager due to two factors: first, the Dance Exchange videos were going to be sent during the same fiscal year to the vendor along with materials for two audio digitization projects. This meant that the SOW had to include the technical specifications for VHS, ¼" open reel audiotape, and wire recordings. Second, because this was the first video mass-digitization project UMD undertook, and the first video project contracted to this vendor, the DCMR manager and the metadata librarian consulted with the vendor for two months on technical specifications, file deliverables, file structure, and metadata deliverables. This necessitated several email and telephone conversations regarding the video technical specifications. For example, standards were determined for the container, extension, bitdepth, chroma subsampling, framerate, timecode, audio channels, and audio quality. However, the pilot project revealed an oversight on UMD's part as the data rate for the derivative file was not determined simply due to a lack of awareness for its necessity. UMD also needed to include details regarding the structuring of the hard drives that the vendor would return. Files require processing in a flat structure for ingest and archiving as complex file structures cause UMD's systems to slow down. The file and technical specifications were determined in part by UMD's established Mac operating system infrastructure and the university's adoption of the Adobe Suite. This necessitated a decision to use .mov uncompressed QuickTime files as the preservation masters, with H.264 MPEG-4 files as derivatives, or access copies. The technical specifications were lengthy enough to warrant inclusion as an attachment to the SOW. The standards set in this SOW have since been utilized for additional small projects digitized by the same vendor.



Digitization workflows and vendor technical specifications (02:46)

The vendor based their cost estimates in the SOW on estimated tape durations as there were few tapes where the length of the content was documented; half (50 items) were estimated at 31-60 minutes, and the other half were estimated at 61-90 minutes. Based on these estimates, it was thought that 6TB of space on external hard drives would be enough for the vendor to deliver the project deliverables. Two unexpected outcomes resulted from these estimates. First, the project cost less than originally estimated because there were fewer tapes at the upper duration end. However, because many of the shorter duration tapes were longer than expected, UMD needed to send the vendor an additional 6TB, though not all the space on all the drives was used.

Without dedicated staff for such projects, people from different departments contributed time to prepare the project for the vendor. SCPA provided staff to pack and prepare the physical media with guidelines from the manager of DCMR. The metadata librarian worked with the collection's curator and the manager of DCMR to map the existing metadata from an in-house inventory into a spreadsheet for ingest and also into a separate spreadsheet for the vendor's shipping manifest. Such a disconnected preparation process is not ideal for projects, whether with a vendor or in-house, and UMD is exploring the feasibility of employing dedicated hourly staffing for similar projects in the future. This disconnected preparation process also caused problems in the project timetable, specifically the shipment schedule, as there were difficulties in metadata mapping, budget and account coordination, and shipment preparation. UMD will use this information to more precisely plan project timelines in the future.

Metadata

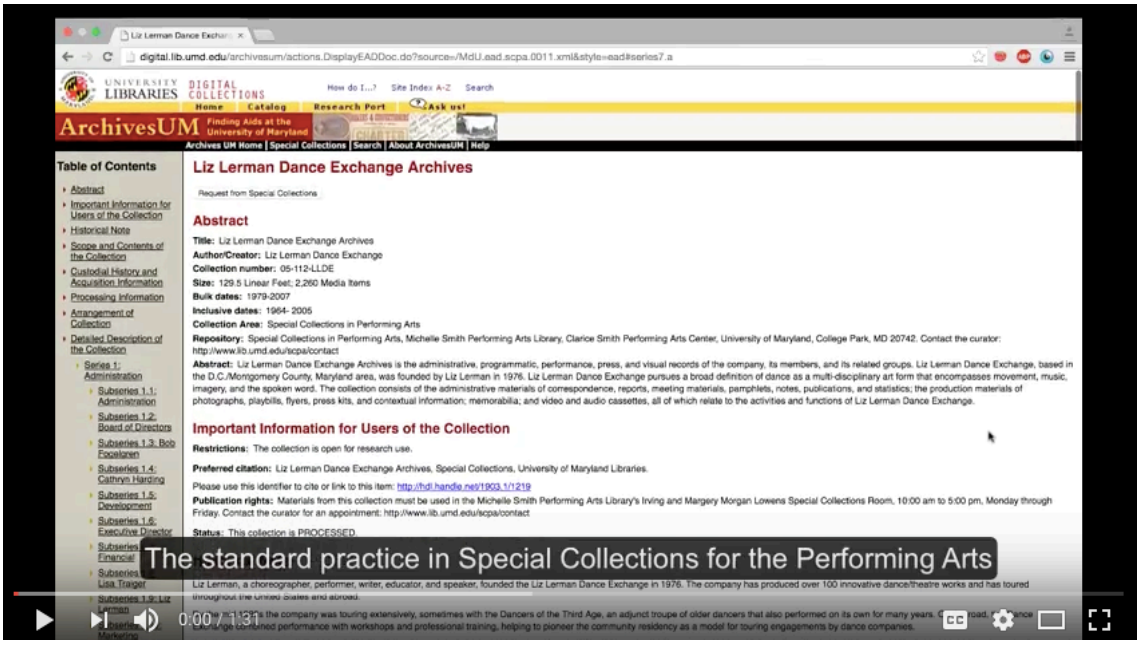
Descriptive metadata

With the vendor SOW finalized, one of the first preparation tasks was to evaluate the current state of the collection's description. Per the SOW, a spreadsheet

with basic metadata formatted to the vendor's specifications was expected as a part of the shipment. The project team also wanted to plan for additional metadata (descriptive, technical, etc.) to collect during the course of the project.

A challenge when planning for the metadata for this project was one that most archival digitization projects face—archival collections are often not described to the item level. Yet, in digitization projects, item-level metadata is frequently expected in order to aid discovery. It is not uncommon to have a simple spreadsheet inventory of archival AV items. Furthermore, archivists commonly have to rely on the handwritten metadata on the objects, which is not necessarily complete or accurate.

All of the above situations applied to this project. The standard practice in SCPA is for collections to be described at the folder level for manuscript materials, with item-level inventories available for select collections. The Dance Exchange collection boasts minimal description at the item level for AV items, which an intern had created from the handwritten or typed labels on the media. This basic inventory was used to complete the metadata for the vendor's template before the shipment. The inventory metadata was also mapped to UMD's local video metadata template, which is used to gather all descriptive, technical, and administrative metadata currently required and allowed by UMD's systems.



Liz Lerman Dance Exchange collection description and metadata (01:31)

Depending solely on the metadata gleaned from the handwritten or typed labels on these VHS tapes proved problematic. During the digitization process, it was discovered that these labels were not entirely reliable. The description supplied on the tapes was not accurate enough to prevent the selection of two duplicate recordings—the vendor fortunately caught the oversight—even with the curator enlisting Dance Exchange staff to assist in identifying priority recordings. This raised the question of what else might be incorrect, but not obviously so. As media collections are digitized

more holistically, it is possible that this is something that can be addressed through increased pre-project appraisal or post-conversion quality assurance (QA) processes.

The inventory available for the collection merely provided a title, a date, and usually a brief description. The minimal descriptive metadata creates a challenge for staff or patrons to successfully retrieve and view the recording for which they are searching. Currently, these videos are only viewed by scholars who are already familiar with Lerman's works. These researchers will contact the company, who will refer them to SCPA. If UMD wishes to expand descriptive metadata, and thus the discovery of these videos, a time-effective way to perform this work must be determined. How do archivists capture and describe details regarding the choreography that may aid scholars? As with all types of materials, generating the level of description desired for discovery may require some familiarity with the subject area, which is not always available. As these videos are about a specific dance company, it may require the expertise of the dance company itself.

Additionally, with time-based media, the only way to know that something is what it purports to be is to view or listen to the item. That is costly in time and other resources, and it is likely not possible to do with the original media due to preservation issues; these preservation problems are one of the reasons digitizing these collections is a priority. As a part of this pilot, the authors are investigating ways to address the metadata challenges that are unique to describing dance archives. Can the Libraries find and fund student assistants with some subject knowledge, such as MFA students, to view the videos in order to add useful description? That is a significant time commitment, so it is vital to investigate the costs and benefits of such an undertaking. Archivists and metadata librarians also need to develop and provide guidelines for description and employ controlled vocabularies to create consistency within dance collections. The authors are investigating using a dance scholarship taxonomy developed by Janet Lansdale and Susan Wiesner.[1] This taxonomy includes controlled terms for genre, style, period, scholarship category, and methodology. Some of these classifications will not apply to the Lerman videos, but will provide a starting point for further work. A dance taxonomy based on this dance scholarship taxonomy was developed for the Resources for College Libraries (RCL). Broad terms are available in the Getty Art & Architecture Thesaurus (dance companies, choreography) and the Library of Congress Subject Headings (types of dances), but these broad terms are likely not granular enough for research use.

Technical metadata

In addition to descriptive metadata, the authors also hoped to collect technical metadata during the digitization process in order to align the digital video practices of the vendor with UMD's existing local practices regarding digital audio. Metadata standards and best practices for digitizing video materials lag behind those for audio and text materials. The first issue that needed attention was file format. With digital audio, the Broadcast Wave Format (BWF) is the standard format for audio preservation

[1] Sudan L. Wiesner and Janet Lansdale, "Analysing the Discourse of Dance Research in Selected Journals 1996-2005" (presented at: Society of Dance History Scholars Conference, Skidmore College, Saratoga Springs, New York, June 12-15, 2008). Research and database development supported by a UVA Libraries Research grant.

master files, with .wav and .mp3 files widely accepted for use copies. A single, widely accepted format does not yet exist for digital video; best practice standards range widely because they are dependent on the software selected and storage capacity of the systems.

When discussing digital video formats, what is actually discussed is the wrapper: a container that includes all the information necessary to play the video. The wrapper can include the separate audio and video streams as well as any metadata and closed captioning files. There is not one single standard wrapper. Rather, institutions usually select a particular video wrapper (e.g. .mov or .avi) based on the operating systems the institutions support.[2] In audio, the preferred preservation file (BWF) is platform agnostic, and thus widely adopted by all regardless of operating system. Because adoption of video format/wrapper is dependent upon the operating system, it is difficult to define a single accepted way to proceed, and one cannot always emulate what another institution is doing. The issue of operating system and software support affected decisions regarding access files. Another way in which digital video lags behind audio is that video wrappers lack a standardized set of embedded metadata fields, such as those found in BWF.[3] Embedding metadata is possible in video, but which fields can be embedded is determined by each video format.

As noted above, due to operating system and software considerations at UMD, .mov was chosen as the wrapper for video preservation master files. In planning for the pilot, the project team wanted to collect embedded metadata similar to that available in BWF in an effort to be consistent among audiovisual preservation masters in UMD's digital repositories. In working with the vendor, it was determined that the technology for embedding that metadata within .mov files is not currently possible. There were issues with including specific fields, either using the Extensible Metadata Platform (XMP) or FFMPEG.[4] FFMPEG was an ideal option, but technologically was not feasible. The vendor found that embedding the desired metadata via FFMPEG rendered the file unplayable. In discussions with the vendor, it was revealed that any metadata the vendor had embedded using either method would disappear when opened in Adobe Premiere.[5] Additionally, the vendor's automation and ingest systems constrained the work, as working with FFMPEG required re-rendering the file, which created insurmountable storage issues due to file size.

At the recommendation of the vendor, UMD instead received a PBCore XML file with metadata that will exist as a sidecar to the preservation master file. PBCore is a metadata standard specific to audiovisual materials, and covers multiple aspects of AV metadata, including metadata about the source item, as well as each additional instance (e.g. preservation master, access copy, etc.). These PBCore sidecar files will serve as

---

[2] Barbara Goldsmith, "Digitizing Video for Long-Term Preservation: An RFP Guide and Template" (New York University Libraries, 2013) accessed September 26, 2016, http://memoriav.ch/wp-content/uploads/2014/07/VARRFP.pdf.

[3] Federal Agencies Audio-Visual Working Group, "Embedding Metadata in Digital Audio Files: Guideline for Federal Agency Use of Broadcast WAVE Files" (2012), accessed September 26, 2016, http://www.digitizationguidelines.gov/audio-visual/documents/Embed_Guideline_20120423.pdf.

[4] More information on these frameworks can be found at http://www.adobe.com/products/xmp.html and https://www.ffmpeg.org/about.html.

[5] Due to the University of Maryland's adoption of Adobe Suite, our quality assurance is performed using Adobe software.

substitute for embedded metadata. These files were not included as deliverables in the original SOW, but the vendor graciously included them in their work and their deliverables to UMD. Consequently, metadata librarians and DCMR will consider including this solution in future video digitization projects.

Quality Control and Quality Assurance

Although the VHS tapes were assessed for signs of degradation six years prior to this project and were found to be in good condition, the vendor experienced playback issues with many of the items, requiring testing on several players prior in order to play them back. Of the 100 tapes sent, 98 were digitized. However, this was fortunately not due to degradation but to the duplication issue noted above. Both the vendor and in-house staff performing quality assurance (QA) noted inherent issues (e.g. beginning cut off, end cut off, audio buzz, audio hiss, audio hum, audio low levels, and video picture breaks), and issues due to degradation impacting 15 percent of the tapes (e.g. audio distortion and loss). A librarian in DCMR completed post-digitization quality assurance. This involved executing automated file checks on 100 percent of the files and performing visual inspection on portions of 25 to 50 percent of the files, randomly selected. Sampled QA was employed as it was not feasible for staff to complete QA for the entire project in the time allotted by the vendor (30 days). QA sampling is used for all vendor-based digitization projects over 25 recordings.

Examples of inherent issues and issues caused by degradation in videotape (04:13)

Access

After ingest into the digital collections repository, the metadata was made searchable and available to all, and the streaming files were set to on-campus access only. Though the Libraries and the donor have a good relationship that has allowed this level of access, it is impossible to grant extended access to the materials without

clearing all rights from every performer, to say nothing of the music licensing considerations. Patrons that cannot come to campus to view the materials can request files for individual use. The Libraries previously established this model for audiovisual digital surrogate access because it preserves the content of the original asset and provides patrons with a surrogate for access when the original is inaccessible due to technological and equipment limitations.

Conclusion

A defining result of this pilot is that it is now possible for UMD to set standard technical specifications and workflows for managing video files, including developing efficient procedures for metadata harvesting, mapping, and enhancement; quality assurance; and ingesting and archiving multi-drive and multi-TB shipments. This process will become the model through which UMD will initiate outsourcing video digitization in the future. Though digitizing additional tapes from the Dance Exchange Archives was not slated for fiscal year 2016 through Libraries funded digitization, the collection's curator is currently working with the company to pursue grant funding for a future digitization project.

This collection warrants such effort and priority, as the Liz Lerman Dance Exchange Archives supports the research needs of dance scholars in many ways. The original sources available in the collection provide unique historical context and insight for creating new research on dance history and performance. While these materials will never take the place of experiencing a dance performance in the present, they do impart a legacy and the means to discover the past. The contents provide one form of performance documentation via photographs, notation, and other paper-based materials, but easily the best manner to experience the work after the fact is through video documentation. This pilot video digitization project will not only serve UMD in comparable future projects, but it brings this collection a step closer to allowing researchers to more fully experience and study the company's performances.

Bria L. Parker is the Metadata Librarian at University of Maryland Libraries, where she creates, transforms, and consults on metadata for a number of projects and collections. She received her Masters of Science in Information from the University of Michigan in 2009. Prior to coming to the University of Maryland she was the Metadata and Digital Collections Librarian at NASA Goddard Library, where she helped launch the Goddard Library Repository. She previously served as the Audio Digitization specialist at the University of Michigan Library, where she co-authored a paper on the challenges of audio for digital collections.

Robin Pike is the Manager, Digital Conversion and Media Reformatting at the University of Maryland, College Park Libraries, managing digitization and digital conversion operations. She previously worked as the Audiovisual Archivist at The Catholic University of America. She holds an MLIS from the University of Pittsburgh. She is an SAA Recorded Sound Roundtable Steering Committee member, and previously served as the chair, and serves on the

Association for Recorded Sound Collections Education and Training Committee.

Vincent J. Novara is Curator for Special Collections in Performing Arts at the Michelle Smith Performing Arts Library, University of Maryland, where he earned his Masters of Music. A Certified Archivist, he has held archivist positions at UMD since 1994, and was appointed curator in 2005. Scholarly communications include book chapters, and articles and reviews in Music Library Association's *Notes*, ACRL's *CHOICE*, and *Educational Media Reviews Online*; as well as conference panelist, presenter, or moderator. In October 2016, Novara offered his twelfth Project Management for the Archival Workplace workshop for the Mid-Atlantic Regional Archives Conference.