Overview

The National Endowment for the Humanities provided $260,000 to the University of Michigan in support of a two-year project (2016-18). The PI is Associate Professor Paul Conway; the co-PI is Professor Kelly Askew. The project's aims are to digitize, preserve, and provide online access to over 900 audio recordings and the associated scripts for *Music Time in Africa*, broadcast by the Voice of America between 1965 and 2005, along with a selection of original field recordings of beautiful and influential African music recorded for the radio programs by musicologist Leo Sarkisian. An online interface will permit exploration of the digital collection and the juxtaposition of audio streaming and script texts. In conjunction with the Voice of America's listener outreach programs, the project will explore open access options and engage performer communities about the release of field recordings and tracks included in the radio broadcasts. The primary outcomes of the project are a model for culturally sensitive online / web-based distribution of digitized musical recordings and a demonstration of the power of community engagement in providing access to scholarly resources on musical heritage.

Major Accomplishments in Year One

1. Set up the project's physical workspace, recruited and hired student workforce.
2. Processed to MediaPreserve all *Music Time in Africa* radio program recordings (1966-2005) found in the Leo Sarkisian Archive that were loaned to the University of Michigan.
3. Completed digitization of these 920 recordings, and completed two phases of quality control on the digital files.
4. Established post-digitization processes, including quality inspection, inventory and files management, preservation of master digital files, and metadata schemes and procedures.
5. Established a project website and made two major presentations on the project at international venues.

Completed Tasks in Year One

1. *Workspace*: The University Library's Special Collections Library provided the project with a flexible workspace in the Buhr Building on south campus. The Buhr Building also houses the University Library's Conservation Lab and the Digital Conversion Unit, providing the project with ready access to advice on preservation and digitization. With the support of Shannon Zachary, head of conservation and preservation, the PI and project staff configured the workspace to support the processing of audio tapes to the digitization vendor. The University Library installed appropriate shelving and furniture and has made available three computer workstations to support the project.

2. *Student staff*: NEH funding provides support for graduate students to work on various aspects of the project. During the first year the PI recruited and trained three UMSI graduate students during the summer 2016 and three students during the 2016-17 academic year:

    a. *Summer 2016:*
        i. Kayla Carucci (tape processing); Leigh Gialanela (finding aid; website development); and
        ii. Jacob Kidd (audio tape processing);
        iii. Stephanie Zang (website development).
    b. *Academic Year 2016-17:*
        i. Jacob Kidd (project coordination; user interface design; access system configuration)
        ii. Leigh Gialanella (files processing and metadata creation)
        iii. Kaitlyn Sisk (metadata creation; OCR experiments)

3. *Package model:* The project team developed a model specifying the elements of a complex package representing the concept of "radio program" from *Music Time in Africa*. The elements include the analog audio recording for a given date and its associated script (if available), along with multiple digital surrogates of audio program and script, all of which are tied together by descriptive and technical metadata. Some elements of the package are preserved in a digital repository, while other elements of the package are integrated into the access system. **Attachment 5** is a summary of the package model.

4. *Finding aid:* The project team created a master inventory of the Leo Sarkisian live field recordings, which are currently accessible only on CDs at the University of Michigan School of Music Library. The foundation for the inventory is an existing tracking spreadsheet created as part of the digitization of 360 field recordings by the University of Michigan between 2010 and 2012. Co-PI Kelly Askew provided detailed supplementary information on each recording, including variant titles and lists of performers. The finding aid will be published through the university's DLXS EAD access platform and will provide a framework for additional access tools for the entire Leo Sarkisian Archive.

5. *Audio tape processing*: The project team identified and processed 920 radio program recordings to the digitization vendor, using shipping protocols established by the vendor. Audio recordings are on either 7" or 10" reels of ¼ inch magnetic tape. The team developed an inventory control and shipping manifest system to support processing of tapes to and from the vendor and follow-up quality control processes. The project team assigned each audio recording one of the following codes, indicating the completeness of the analog components of the package:

    a. **Platinum**: an audio recording of a full 25 to 30 minute radio broadcast [including the voice of the announcer] and the associated script.
    b. **Gold**: an audio recording of a full radio broadcast without the script.
    c. **Silver**: the musical inserts for a radio broadcast and the associated script [missing the voice of the announcer]
    d. **Bronze**: incomplete broadcasts or scripts.

Attachment 1 is a graphic representation of the distribution of the programs in terms of completeness across program dates. The curve in the graphic reflects the tendency to cannibalize earlier program content for later broadcast and the prevalence of re-broadcasting of previous programs at later dates. The final performance report on the project will provide a more detailed analysis of the distribution of program recordings and scripts over time.

6. *Audio recording digitization:* The project team processed 920 recordings to the third party vendor. The recordings were processed in batches of 100 tapes. Each physical recording is associated with a unique University of Michigan Library barcode, which the vendor used for inventory control and the project team uses for file management, including naming conventions in the preservation repository. By the end of the first year (31 May 2017), all recordings sent to the vendor have been digitized, the digital files have been received, and vendor compliance processes have been completed on the digital products. Additionally, the original source tapes are housed in the library's preservation facility, the archival master recordings are ingested into the university's preservation repository, and the contract with the vendor has been fulfilled.

Digitization Vendor

The MediaPreserve, an audiovisual laboratory in Cranberry Township, Pennsylvania, is the vendor selected to digitize the 900+ radio program tapes. MediaPreserve is well equipped, highly experienced, and fully capable of completing the digitization processing on the schedule required for the grant. MediaPreserve provides: customized services needed to obtain preservation quality master files; production files identical to master files except for normalization and noise reduction processing; and appropriate access derivatives.

Figure 1 is a table of the structure of the digital files delivered by the vendor for each digitized radio program. The files include an archival master (WAV), normalized access master (WAV), streaming file (MP3), bitmap images of the reel box cover and magnetic tape reel (JPG), and checksum calculations for the preservation, access, and streaming files (MD5).

```
Directory of F:\Audio_Deliverables\39015098237145

10/26/2016  11:37 AM    <DIR>          .
10/26/2016  11:37 AM    <DIR>          ..
10/25/2016  03:09 PM            38,819 39015098237145.xml
09/29/2016  04:40 PM       465,435,184 39015098237145_access.wav
09/30/2016  12:56 AM                62 39015098237145_access.wav.md5
09/14/2016  10:59 AM         2,051,448 39015098237145_BoxSide.JPG
09/29/2016  04:39 PM       465,435,182 39015098237145_preservation.wav
09/30/2016  12:56 AM                68 39015098237145_preservation.wav.md5
09/14/2016  10:59 AM         1,771,686 39015098237145_ReelFront.JPG
09/29/2016  04:43 PM        38,786,821 39015098237145_streaming.mp3
09/30/2016  12:56 AM                65 39015098237145_streaming.mp3.md5
               9 File(s)    973,519,335 bytes
```

Figure 1. Table of digital files delivered for each radio program recording

Preservation master and normalized (access) files conform to the highest standards recommended by IASA TC-04, the accepted international standard for audio digitization. In this project, MediaPreserve provided the technical metadata for each digital file, with all files for a given batch wrapped in a master METS record that then receives further descriptive and administrative metadata as part of quality review and post processing activities. For more information about MediaPreserve, see: http://ptlp.com/en/mediapreserve/overview/about-us/

7. *Preservation of archival master digital files:* All of the digital files received from the vendor have been uploaded into the University of Michigan's "dimly lit" digital preservation repository. Each file includes a checksum to facilitate the periodic monitoring of the files for possible corruption. For more information on MiStorage services, see: http://services.it.umich.edu/mistorage

8. *Metadata model:* The project team developed, tested and implemented a customized metadata model for the combination of radio program recording and associated script. The model is an adaptation of the Metadata Object Description Schema, which is a widely recognized metadata standard maintained by the Library of Congress. The MODS record for each radio program contains the following twelve data fields, all of which conform to the MODS standard for format and content.

   a. Title   [Music Time in Africa date]
   b. Proper Names   [all named individuals, plus program director and host]
   c. Role(s) (of person(s) named) [e.g., host, performer, producer]
   d. Type of Resource   [audio recording]
   e. Genre   [script]
   f. Publisher Name   [Voice of America]
   g. Place of Publication   [Washington, DC}
   h. Publication Date   [date of radio program broadcast]
   i. Capture Date   [digital capture: 2016 or 2017]
   j. Languages   [English + other languages identified in script]
   k. Subject   [Geographic coding]
   l. Note
      i. OCR text (script/audio)   [extracted from ResCarta file]
      ii. Duration   [hh:mm:ss]
      iii. Completeness of program   [recording + script]
      iv. Notes   [anomalies]
      v. Barcode control number   [link to source media and digital files]

The project is making optimal use of the open-source ResCarta Toolkit to produce radio program transcripts, descriptive and technical metadata, and appropriate derivatives. Information on the ResCarta Toolkit can be found here: http://www.rescarta.org/. In October 2016, the head of the ResCarta Foundation, John Sarnowski, visited with the project team, provided specialized training on the Toolkit for project staff, and presented a workshop/lecture in a graduate course at the University of Michigan School of Information.

The grant proposal had suggested that the project would adopt the PBCore metadata schema. After completing a detailed comparison of MODS and PBCore, the project team decided that the MODS model was most appropriate for the material in hand; in addition, the MODS model was a much closer fit to the metadata scheme of the access delivery system supported by the University of Michigan (see below). Utilizing the PBCore scheme would have required custom programming in the access system. **Attachment 2** displays version 1.1 of the metadata fields that are likely to be displayed to the end user. The final performance report will provide more detail on the metadata structure in the access system, including field that are and are not displayed to the end-user.

The final performance report will include more detail on the metadata scheme, references to official MODS documentation, and information on cataloging procedures. Some of this information will be posted to the project website prior to the completion of the project.

9. *OCR testing and evaluation:* The *Music Time in Africa* scripts and the full audio programs provide a wealth of valuable metadata and information that we want to extract and use to describe the program and support end-user search and discovery. The project team conducted systematic experiments to determine the extent to which the printed script or the audio program could

provide the best source of a text transcript (and also to understand the complexities of correcting "dirty OCR" from the audio transcription process). **Attachment 3** is an executive summary of a more comprehensive report that will be published to the project website. The bottom line: create searchable text from the digitized script when such a script exists; digitize the script at 300 dpi/24-bit color for best results; digitize from the audio recording where one a full program exists and where the script is wanting (cleanup errors to support search).

10. *Testbed of 31 programs:* The project team completed initial processing on a selection of 31 programs chosen by co-PI Kelly Askew from across the spectrum of programs in the collection. Prof. Askew applied her in depth knowledge of African heritage music. The 31 programs represent a test bed for working through metadata procedures, perfecting the metadata scheme in the access system, and configuring the user interface. The testbed is now live in the Kaltura streaming service, but with access restricted to project personnel, pending additional work on intellectual property rights.

11. *Wire framing a user interface:* A *Music Time in Africa* radio program is a complex object that exists as an international broadcast at a single point in time. The major components of an ideal representation of a program are: its date of broadcast; a script; a recording of the program; and information about the content of the program (metadata). In sub-optimal circumstances, a script may be missing, the recording consists of only the musical inserts without the voice of the announcer, or missing components of either script or recording. An important goal of the project is to represent together the components of a radio program, whether ideal or sub-optimal. The project team put considerable effort into modeling a user interface to meet the goal, including evaluating other digital projects that incorporate text and audio recordings, exploring the potential and the limitations of the university's access systems, and investigating the capabilities of the ResCarta Foundation's java script-based web gateway tool. **Attachment 4** is a preliminary concept wireframe of a possible custom programmed user interface to *Music Time in Africa* radio programs. In the second year of the project, the team will complete the development of a user interface that can connect to the university's enterprise-grade audio-video streaming service. This approach ensures that access is persistent, even if the user experience changes over time.

Year One Reporting and Presentations

1. *Project Website*: The project team developed an informational website for the project, built on the University hosted WordPress platform. The URL for the project website is: http://mtia.sites.uofmhosting.net/. The project website is distinct from the content management system that will deliver digitized content and manage access protocols for the system. The final performance report will provide the specifications for the access system.

2. *Society for Ethnomusicology Annual Meeting, November 12, 2017, Washington, DC:* Kelly Askew and Paul Conway assembled a panel at SEM focusing on the legacy of Leo Sarkisian. Conway presented their joint paper on the digitization *of Music Time in Africa*. Current host Heather Maxwell contextualized the project in the context of the long history of Voice of America's broadcasting to Africa. NEH program officer Jesse Johnston introduced the panel.

3. *King's College London, Department of Digital Humanities, May 16, 2017*: Paul Conway delivered an invitational talk on the project's emergent approach to access to radio programs that contain original ethnographic field recordings. Panelists Karen Colbron (Jisc) and Bill Thompson (BBC) commented on the wider implications of the paper on access to the audiovisual heritage resources in Africa.

Goals and Activities for Year Two

After one year, the project is on target with its workflow and positioned for a very successful second year. Among the many tasks needed to meet all of the specified goals, the following are the most significant.

1. *Project Advisory Board*: Convene the members of the external Advisory Board in Ann Arbor and seek specific advice regarding access issues, access system design, and promoting the work of the project to wider audiences.
2. *Process the radio programs:* Complete the digitization of extant program scripts, complete OCR processing on all digitized scripts, create and clean up audio transcripts for those programs missing scripts; complete metadata generation for all radio programs in the collection (ca. 900+); proof and correct any metadata errors.
3. *Populate the access system:* Ingest into the University's Kaltura system (MiVideo) the streaming access version, the digitized scripts, and the associated MODS metadata record for 900 radio programs. Utilizing Kaltura's very sophisticated intellectual property management module, establish limited access to the 900 radio programs, initially to University of Michigan authorized users.
4. *Create public-facing web gateway*: Develop custom java-script programming for public-facing interface that supports the responsive juxtaposition of digitized scripts, radio programs, descriptive metadata, and historical images as well as faceted browsing of radio program content.
5. *Explore Ethical Access options:* We aspire to make as much of the radio programs and associated Leo Sarkisian field recordings available to as wide an international audience as possible. The PI and co-PI will lead a thorough review of access options, working closely with the project's Advisory Board and with the legal team at the University of Michigan.
6. Outreach to African communities: The PI and co-PI will develop one or more small scale pilot efforts to connect the Sarkisian field recordings and associated field recordings to communities from which the music originated, focusing initially on the music of Ghana, Liberia, and Tanzania. The pilots will be conducted against a backdrop of sensitivity to the many complexities of access to heritage music in local communities.
7. *Presentations, papers, reports:* Kelly Askew and Paul Conway will present a paper at the International Association of Sound Archives (IASA) annual meeting in Berlin (Germany) in September 2017. While in Berlin, Kelly and Paul will consult on the project with the Berlin Phonogramm-Archiv. Kelly and Paul will revise and submit the paper read at King's College London. Additionally, we are planning additional research papers on the project in the coming year, particularly focused on varying perspectives on intangible cultural heritage.
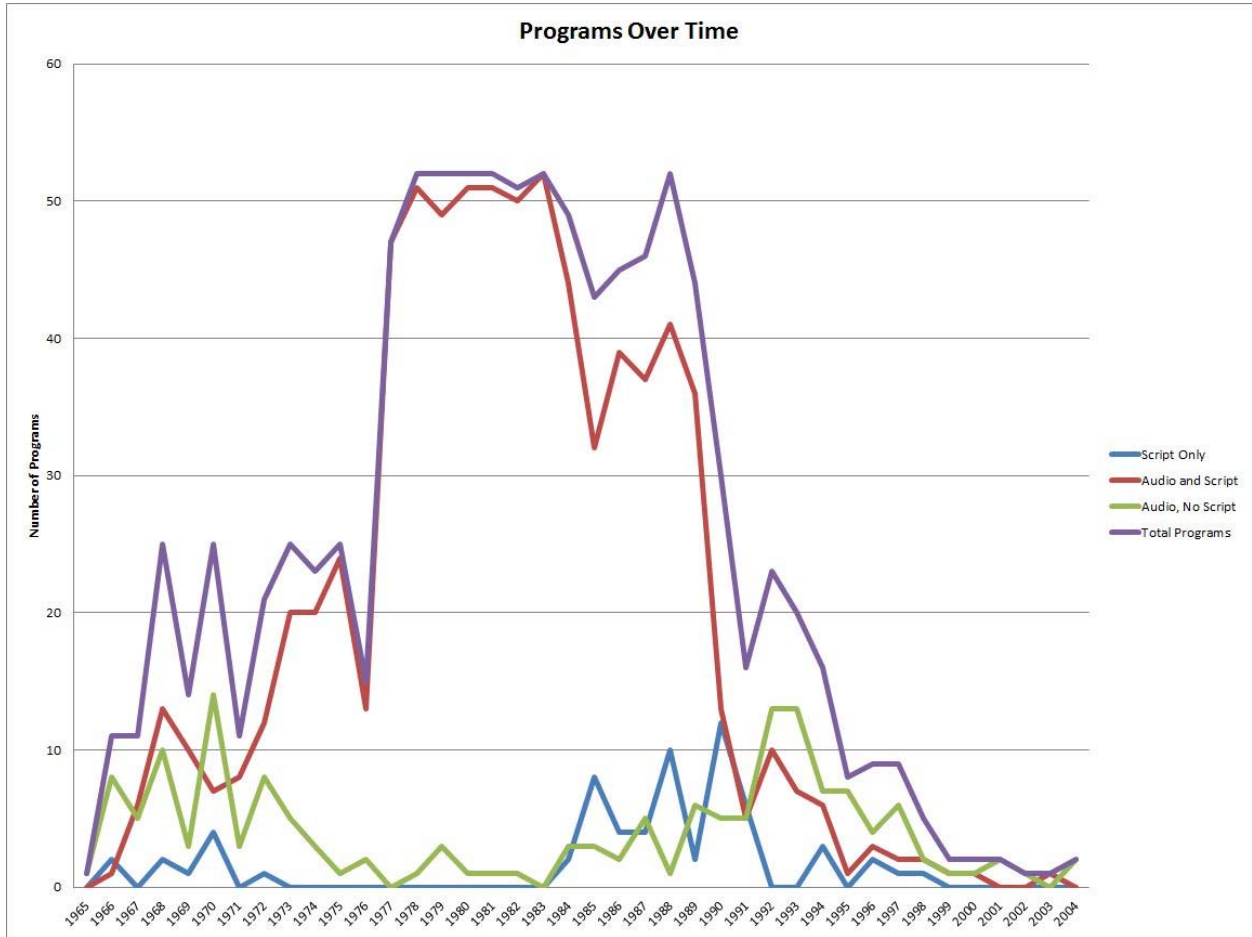
Respectfully submitted,

Paul Conway, Associate Professor
Principal Investigator

16 June 2017

Attachment 1. Distribution of Radio Programs by Completeness Across Time, 1965-2005

Attachment 2. Metadata Fields Displayed in Kaltura Public Interface

MODS Name	Kaltura Display Name	Type	Controlled Values	Notes
Title	Program Name	String		
Role/Name	Host	String		Pull host name from the role "host"
Publisher Name/Place of Publication	Origin	String		Combine the Publisher Name and Place of Publication for this display, separated by a comma
Capture Date	Digital Reproduction Date	Date		
Note: Completeness	Completeness	String (potentially number in the future)	Platinum, Gold, Silver, Bronze	Change the display based on the value: If platinum, display "full radio program + script" If gold, display "full radio program" If silver, display "program inserts + scripts" If bronze, display "incomplete"
Size (inches): digital origin	Script Dimensions	String		In inches
Part-extent: end	Script Pages	Integer		
Note: Duration	Program Duration	String		Minutes and seconds
Subjects: Country	Country	String		Repeatable (will have more than one country listed)
Note: Notes	Notes	String		

Attachment 3. OCR Testing-Executive Summary

Ethical Access to *Music Time in Africa*
NEH – PW-234763-16

Optical Character Recognition of Program Script Content
EXECUTIVE SUMMARY

Summary

The *Music Time in Africa* scripts and the full audio programs provide a wealth of valuable metadata and information that we want to extract and use to describe the program and support end-user search and discovery. The audio transcription and optical character recognition tools in the ResCarta Toolkit can extract this information, but both require at least some time to clean up. Since a large number of the programs are full broadcasts with scripts, we initially wanted to know whether the audio recording transcription or the optical character recognition (OCR) of the typed script produced the most accurate transcript. We learned that the OCR transcripts were far more accurate than the audio transcriptions. Then, we wanted to find the optimal digitization settings to produce the most accurate OCR conversion of the script, using an untrained Tesseract OCR engine included with the ResCarta Toolkit. In a sample of five typical scripts, we found that digitization at 300 DPI in 24-bit color produces the cleanest OCR transcripts. A major implication of these results is that our workflow can be more efficient; we can use one set of scans for OCR and for images displayed in the access system.

OCR Versus Audio Transcript

Five scripts from 1979 were digitized at 300 DPI in color and run through the OCR program, and the associated audio programs were run through the audio transcript program within ResCarta. We determined and highlighted important words (any words that had significant information or value that needed to be accurately captured) on photocopies of each script. These important words were checked for accuracy in the OCR transcript and the audio transcript.

Table 1 compares the rate of error of all identified important words between the OCR results and the audio transcription results. The number of important words that contained errors are listed under "# important words wrong" for both the script and the access audio files. The percentage of important words with errors out of the total number of important words is shown under "% wrong" for both the scripts and the access audio files. The average of each column is displayed in the last row.

From table 1, we can see that the OCR tool is more accurate than the audio transcription tool for each of the five scripts. On average, the audio transcription produced an important word with an error more than twice as often. The audio transcription program does not have the capabilities to recognize most of the important African words, other than country names. Even then, the audio transcription was not completely accurate with countries.

	Script (text)		Audio (recording)		
Program Date	**# important words wrong**	**% wrong**	**# important words wrong**	**% wrong**	**Total number of important words**
April 29, 1979	33	21.02%	103	65.61%	157
May 6, 1979	33	23.24%	102	71.83%	142
May 13, 1979	81	52.94%	108	70.59%	153
May 20, 1979	72	42.86%	112	66.67%	168
May 27, 1979	17	11.04%	119	77.27%	154
Average	**47.20**	**30.22%**	**108.80**	**70.39%**	**154.80**

Table 1. Comparison of error between script OCR and access audio transcript

OCR Digitization Settings

Once we determined that the OCR was more accurate than the audio transcription, we wanted to know what digitization settings produced the most accurate OCR transcript. We sent a number of test scripts to the Michigan Library for OCR transcription with their OCR program, and discovered that ResCarta produced more accurate transcripts. We then digitized three test scripts (a pink dot matrix script, a white dot matrix script, and an original photocopy script) at 300 DPI, 400 DPI, and 600 DPI in color, and then converted each scan to bitonal for a total of six test instances for each script. Each test instance was run through ResCarta, and a small portion of less than 100 words from each transcript was checked for accuracy.

The table below shows the number of incorrect words and the percentage of words incorrect from each DPI setting in bitonal and color for each of the three scripts. The total number of words checked in each script is listed in the rightmost column.

		24 bit Color		1 bit Bitonal		Total
Script	**DPI**	**# words incorrect**	**% words incorrect**	**# words incorrect**	**% words incorrect**	**Total words**
White	300	**2**	**2.56%**	16	20.51%	78
	400	8	10.26%	43	55.13%	
	600	31	39.74%	73	93.59%	

Pink	300	**17**	**19.32%**	51	57.95%	88
	400	36	40.19%	68	77.27%	
	600	66	75.00%	88	100%	
Original	300	**8**	**9.76%**	15	18.29%	82
	400	13	15.85%	11	13.41%	
	600	22	26.83%	19	23.17%	

Table 2. Results of OCR test for pink and white dot matrix scripts.

Scripts were scanned at 300 DPI, 400 DPI, and 600 DPI in color and bitonal. Most accurate transcripts for each script are bolded in the table. Table 2 shows that the 300 DPI color scans of both the white and pink scripts were the most accurate for each set of test words. The color scans were more accurate than the bitonal scans at every DPI setting. Increased resolution resulted in less accurate transcripts for both color and bitonal scripts. This study revealed that a 300 DPI color scan of scripts will be sufficient both for optimal OCR transcripts and for images in the access system.

Kaitlyn Sisk
University of Michigan School of Information
24 March 2017

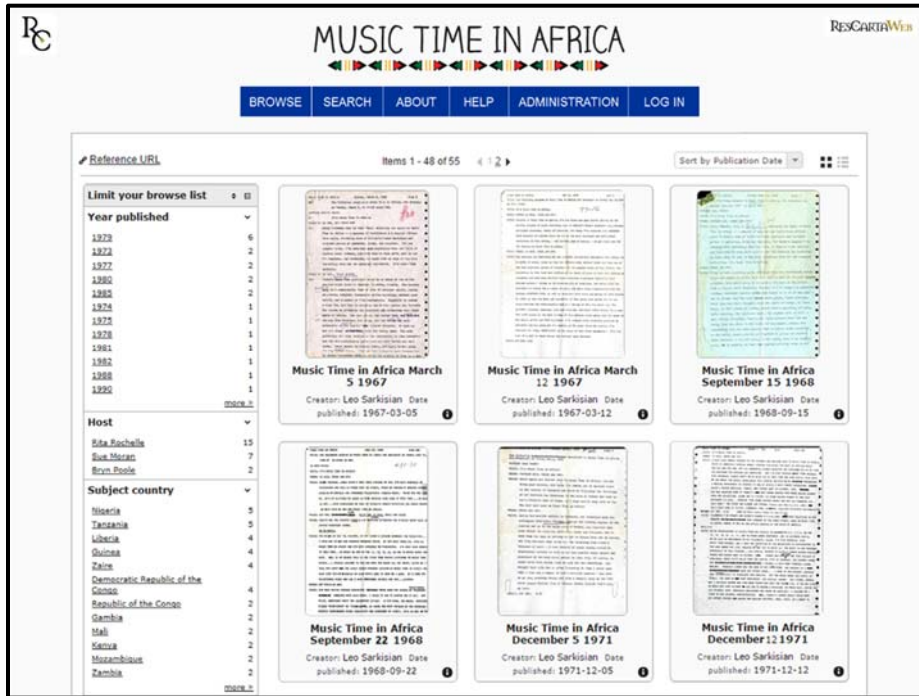**Attachment 4. Preliminary Wireframes for a Possible User Interface.**

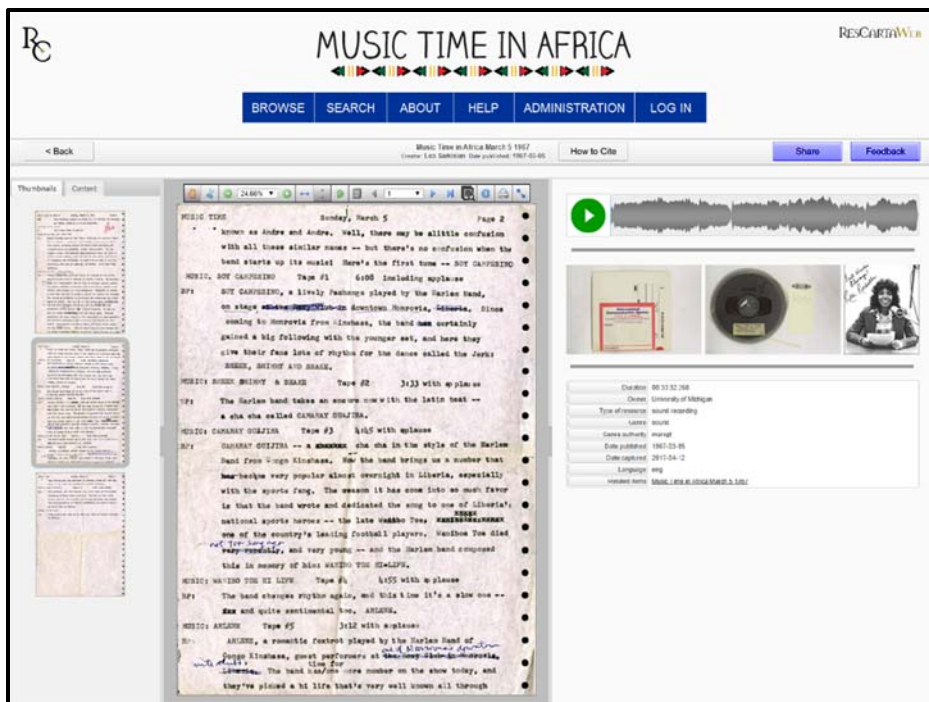**Figure 1. Preliminary wireframe for search results**

Figure 2. Preliminary wireframe for displaying script, music streamer, and metadata

Attachment 5. MTIA Package Elements [ver 1.4, 3/31/17]

This model encompasses the physical objects of recording and script, as well as a variety of associated digital surrogates, that comprise a complex object that is a MTiA radio program. Those elements marked with "A" indicate the elements that will likely be part of a public access system. The proposed disposition of each element is indicated in the right column; some locations have yet to be determined. For each element, additional technical and administrative documentation will be prepared.

		Access	**Disposition**
Sound Recording			
Tape recording – best version for a given date	(physical artifact)		Library storage
• Platinum, Gold, Silver, Bronze			
• Finding aid may note duplicates and alternative versions			
Master file (BWF)	(IASA standard)		MiStorage
Normalized file (BWF)	(IASA standard)		MiStorage
Access file (MP3)		A	MiVideo / MiStorage
Alternative derivative (OGG)		A	
Program Script			
Source script (2-5 pages)	(physical artifact)		Library storage
• Paper typescript (variously legible)			
• Adobe PDF copy from VoA (in absence of paper version)			
Digital image surrogate of script	(.tiff or.jpg)		DLXS / archive
Display derivatives of script	(.jpg / .pdf)	A	tbd
Thumbnail image (on the fly)	(.png)	A	tbd
Full text of script (one or more of following):	(.txt)		
• Raw unformatted (from ResCarta/sound recording or script			UMSI R:
• Manually corrected (from ResCarta OCR)		A	MiVideo
Formatted display version	(.txt / html)	A	tbd
Metadata			
Unique ID for package and its elements		A	MiVideo/DLXS
EAD finding aid			DLXS
IP status indicators (open, snippet, closed)		A	MiVideo
METS File (or other metadata scheme)			
Descriptive (MODS compliant)		A	MiVideo
Technical (MIX compliant)			tbd
Administrative (other than IP status)			
• Preservation metadata			PREMIS
Structural (file structure)		A	tbd
Optional Digital Elements			
Visual images			
• Tape box cover/tape reel	(.jpg / .png)	A	MiStorage
• Archival images	(.jpg / .png)	A	tbd