

Digitization in the Real World

**Lessons Learned from
Small and Medium-Sized
Digitization Projects**

Edited by
Kwong Bor Ng & Jason Kucsma



Metropolitan New York Library Council

Published in the United States of America by
Metropolitan New York Library Council
57 East 11th Street, 4th floor
New York, NY 10003-4605
p: (212) 228-2320 f: (212) 228-2598
Web site: <http://www.metro.org>

ISBN: 978-0-615-37995-2

Cover Design: Jason Kucsma (*illustration by Smartone Design,
licensed via iStockphoto.com*)

Reviewers Committee: Mark F. Anderson, Jill Annitto, Anna Craft, Jody DeRidder, Renate Evers, Wei Fang, Maureen M. Knapp, Sue Kunda, Mandy Mastrovita, Ken Middleton, Emily Pfothenauer, Mark Phillipson, Alice Platt, Mary Z. Rose, Stacy Schiff, Jennifer Weintraub, Andrew Weiss.

Copyright © 2010 by Metropolitan New York Library Council. No part of this book may be reproduced in any form or by any means, electronic or mechanical, including photocopying, without written permission from the publisher.

The views expressed in this book are those of the authors, but not necessarily those of the publisher.

Managing Rights in a Medium Scale Audio Digitization Project

Barbara Taranto and Elizabeth Bradley
(New York Public Library)

Abstract

In 2007 New York Public Library received a small grant to digitize a selection of audio recordings of public programs, including lectures, interviews, and panel discussions that took place at the Humanities and Social Sciences Library between 1983 and 2003. This chapter briefly discusses the scope of the project and the institutional reasons for the initiative. The chapter then discusses some of the significant challenges regarding Intellectual Property Rights, faced by the Library and the strategies it employed to deal with the issues. The chapter concludes with a discussion of the policies and procedures the Library has since put in place to manage audio rights.

Keywords: Access, Audio, Copyright, Media, NYPL, Privacy, Public library, Recording, Taping, Transcription.

The Genesis of the Project

New York Public Library has been engaged in digital projects since the mid 1990's when it launched its first public website. In 1998 it published a compilation of digital "reprints" African American Women Writers of the 19th Century (*Webpage Dodson*, 1998) of out of print, and essentially unavailable 19th century materials from the holdings of the Schomburg Center For Research In Black Culture. By 2000 the Library had established a new unit, the Digital Library

Program, to manage the creation and publication of digitized content from its collections. Within the year an even larger project – the digitization of half a million pictorial items – NYPL Digital Gallery (see *New York Public Library Digital Gallery*, n.d.) - was undertaken and by 2003 the Library was producing approximately 1250 new image files – complete with metadata - per week.

Public domain items were chosen for the first digitization efforts. This allowed the Library to leverage existing bibliographic records for metadata creation without additional research into intellectual property rights. It also allowed the Library to provide a new and exciting online resource for the public.

As the program matured policies and best practices were developed, including how to manage the rare but occasional issue of Intellectual Property (IP). The metadata system was modified to record Intellectual Property rights and access permissions. Processes were established to redact items that were mistakenly digitized and/or published before IP permissions had been received, and new content was being created daily. At the same time the Library began experimenting with new formats such as geospatial data and media files. It also began to look closely at user needs and at new avenues for distribution such as mobile and cellular applications.

As patrons became more experienced web users and their home computing technology improved, user expectations for easy and quick access to rich content increased substantially. Public domain materials were no longer sufficient. Consequently in 2007, the Library decided to initiate a pilot project to determine what was involved in managing rights-encumbered collections in the digital environment. It requested and received a grant from the Metropolitan Library Association to digitize recordings of well-known public personalities and ultimately make them available to the public via the web under the title *24 Hours at NYPL*.

The Scope of the Project

We wanted to test the waters for audio rights management at an institution with a mandate for accessible content delivered at no

charge. There were three key goals: 1) discover the issues involved in rights encumbered audio; 2) determine best practices for resolving these issues; and 3) develop a workflow to manage collections with similar issues. A small sample from a contemporary, but relatively high profile collection would be suitable for this purpose since the subjects involved were already rights-savvy performers. It was also decided that to the best of its ability, the Library would seek permission from these rights holders to make these materials freely available to the public without cumbersome access controls.

The Literature

There is a plethora of documentation on Digital Rights Management (DRM) and DRM products available to leverage and control access to digital assets for commercial gain and/or Intellectual Property management (e.g., ARALOC, 2010, Stream Media, 2010 and Discretix, 2010.) Everything from watermarking to anti-cloning technologies is available to “lock-down” content. Likewise there are as many Digital Library resource pages (e.g., Berkeley Digital Library, 2007) detailing these technologies and the consequent policy implications for managing copyrighted and licensed media in a library setting. Furthermore, there is a very lively scholarly discourse in the preservation community around the proper expression (language) of digital rights and the sharing of digital rights information (see Premis, 2010). All of which is exceedingly helpful once the data have been obtained.

There are many practices employed when engaged in digital activities such as national standards for reformatting, file naming conventions, persistence, etc (e.g., Washington State Library, n.d., and North Carolina ECHO, n.d.). However, the process and procedures by which one should manage a Digital Rights project in a library – especially a public library - setting are not thick on the ground. In fact, it is for this reason that New York Public chose to work on these collections – viz. this is relatively new ground.

We chose to digitize a selection of audio recordings of public programs, including lectures, interviews, and panel discussions that

took place at the Humanities and Social Sciences Library between 1983 and 2003. The recordings, holdings from the Public Education Program (PEP) represent some of the best of the thousands of public programs offered for a small fee to all comers by the Humanities and Social Sciences Library during those decades. Many of these events were held in a grandly restored space called the Celeste Bartos Forum, a venue that has, as a result, become familiar and beloved to generations of New York City culture-seekers.

The programs chosen for digitization included talks and interviews featuring world-renowned writers, thinkers, and public intellectuals such as Chinua Achebe, A.S. Byatt, Umberto Eco, Christopher Hitchens, Robert Thurman, Oliver Sacks, Esmerelda Santiago, Sarah Vowell, and Wendy Wasserstein, among many others. Permission to record the programs was sought from each subject prior to the event. They were originally recorded live, in analog formats including reel-to-reel and cassette tapes.

When choosing these particular programs, preference was given to those subjects whose dynamic performances were the most in keeping with the goals of the digitization project, and for which permission was anticipated to be negotiated and obtained without too much difficulty or searching. For these reasons as well, single subjects (interviewees) were preferred

Rights Issues with Recorded Programs

For the most part flat art works (as plates or reproductions or original art works) are associated with a wide but manageable universe of rights holders. Even maps which have more “creators” than most print materials – engravers, cartographers, drafters, navigators, colorists, etc. – are fixed. The rights pertain to certain aspects of the creation of the object but do not extend to the subject of the work such as “the State of California” or to all parties that participated in the making of the work such as the oiler who serviced the printing press.

The digitally reformatting of the audiotapes was unexceptional. The tapes files were sent to a reputable vendor and quality control was performed on the files when they were returned to the Library.

Technical metadata was recorded and deposited into a database designed for keeping preservation technical data. The Library's metadata specialists using the in-house utility created descriptive metadata. However, obtaining permission from the speakers on these recordings proved to be an unexpected and ongoing challenge.

The first step in rights process is to determine who is the rights holder. This involved contacting the subjects and in some cases editors, managers and lawyers, to match the audio tapes to their rights holders. While the subject was clearly known, the rights documentation needed to be created (with the assistance of the NYPL General Counsel's Office) and managed by NYPL staff, in addition to the management of the media asset. In this case, the Library deliberately chose items where the rights were clearly held by a known body, preferably one with a cordial, ongoing relationship to or affinity for the mission and work of the NYPL. Nonetheless seeking and obtaining permission even when the rights holder is undisputed is often tedious and frustrating. In fact, a good deal of time was spent finding and attempting to contact the subject in question. Once contacted, there were many issues to resolve.

Rights holders were often difficult to reach. Sometimes this was due to the age of the information (addresses are out of date, names have changed, etc) or the amount of time that had elapsed since the creation of the content. Sometimes this was due to privacy needs or the sheer number of intermediaries (such as editors, agents, and business managers) that had to be included in the negotiation process. Sometimes there was no one left to contact. The rights may have reverted the estate in the case of an individual's death or passed to another individual without a clear record. In these cases it was nearly impossible to determine the rights.

In a few cases permission was obtained when considerable time was invested. Often rights holders were unfamiliar with the digital environment and needed to be educated by Library staff about the nature and purpose of digitization. This was a slow process and took a good deal of time and patience and to explain everything in adequate detail. It also required a finesse to describe technical processes in a

way that was meaningful to non-technical individuals. Sometimes this was successful, sometimes not.

Certain more commercially successful individuals, whose fame had increased in the ensuing years, retained counsel that aggressively challenged the Library's request – sometimes with and sometimes without cause (e.g., An individual may not want to relinquish rights to an unedited version of a performance to avoid the possibility of incurring liability for comments made during the performance.) These individuals seemed to be motivated by one of two factors – control over the distribution of creative content produced at an earlier stage in the performer's career, and/or the possibility of generating revenue from the digitized program.

Certain rights holders were unable to remember the contents of the programs they had recorded. Others were uncomfortable with the idea that their material would be available for free streaming broadcast and download via iTunes University. Despite iTunes' seemingly universal brand, the prospect of an NYPL partnership with a commercial vendor confused and possibly alienated some of the performers (or, as was more often the case, their legal counsel).

One strategic decision made by Library staff was to commission written transcripts of the audio recordings of the events that we wished to digitize, so that we might make the scripts of the events available to anxious authors, publishers and literary agents, many of whom suggested that they could not remember the content of a talk delivered so many years before, and were therefore reluctant to give their approval in advance of reviewing it.

The issues presented as follows:

- Rights holder in dispute
- Rights holder incommunicato
- Rights holder is the estate of a deceased person
- Rights holder unfamiliar with technology
- Rights holder represented by aggressive counsel
- Rights holder unable to remember contents of discussion

- Rights holder uncomfortable with venue

For the most part subjects who participated in these public programs were initially hesitant to grant permission to digitize and republish the program on the web, which for many of them constituted a new and unusual forum for the distribution of their work. Interestingly, younger performers gave permission with considerably more alacrity and ease than did older performers (or their counsel). As a result of these unexpected challenges, only a limited number of the digitized public programs have been made available to the public. However, the process of negotiating for rights and gathering permissions is still ongoing, and over time many more subjects have been persuaded to participate in the project.

The digitized materials are preserved for posterity and the existence of this trove of live audio material featuring celebrated authors and thinkers (all of whom have book holdings at the Library, and some of whom have deposited their personal papers or manuscripts, as well) is a source of great interest and enthusiasm for Library staff and the users who have learned about them, thus far.

Lessons Learned

Few of the digitized tapes have been made available for public access. However, the project met two of the three key goals: the issues involved in rights encumbered audio were discovered and a workflow for handling similar issues was established. The third goal - establishing best practices for obtaining permission - can be derived from the case studies in the project.

Many of the subjects were initially reluctant to grant permission for very specific reasons. It leads to the conclusion that addressing these specific concerns can lead to more successful negotiations. By educating the rights holders (and sometimes the counsel) about the new medium of digital audio, many of their material concerns were allayed, or eliminated. By providing written transcriptions as memory aids and evidence to be inspected by counsel, some subjects were able to make a positive, informed decision regarding the Library's request. By being sensitive to privacy issues, ambassadorial about digital

content, and above all else, being politely persistent, the Library has been increasingly successful.

New Practices Born From *24 Hours at NYPL*

The entire process of the project is dependent on documentation. Without the proper releases in place these projects are not possible. If releases do not exist then they must be obtained and rights must be clearly spelled out. Furthermore, all records should be managed in a centralized repository so that future uses for these items are possible. The library was able to identify individual rights holders by sheer tenacity and willingness to put in the hours required. The procedure is prolonged and often takes many turns. Clear guidelines and processes are needed in order to track progress and to properly manage the relations established during the negotiations. Without proper record keeping the process breaks down and accuracy suffers. Because of excellent record keeping the Library was able to pursue this project and obtain as many permissions as it has.

The challenges encountered in obtaining rights for *24 Hours at NYPL* highlighted the need for the Library to develop a more robust rights digital management system. While releases had been obtained for the original recordings these releases were paper documents kept in separate files. There was no database to manage these documents and or the relationship the documents had to the assets. All this work had to be done before the work of seeking permission could begin. Furthermore, evaluating the releases to determine the extent of the Library's rights was necessary before any approach could be made.

These preparatory tasks are often considered incidental to digital projects since digital rights are often discussed as a separate case or special case. However, *24 Hours at NYPL* underscored the need to tie digital rights management to general collection management. In this case, the relationship was drawn between the previously obtained releases and the audio assets. In the case of the earlier project *Digital Gallery* the issue was obscured because of the decision to include only items that were either in the public domain or obtained for the specific purpose of being included in the gallery. Nonetheless, the accurate

documentation in the bibliographic record was the basis for the management of the digital rights.

This understanding has led the Library to create a robust Digital Rights Application that will manage scanned paper documents such as Deeds of Gift as well as rights information. The data will be managed centrally and tied to both the bibliographic databases and the Fedora Digital Repository. Rights data that is stored in a structured format can be shared with multiple systems and exported for administrative purposes and exposed for OAI harvesting.

Additionally work is being done to educate staff to consider digital rights information as vital as provenance to the acquisition process; and to include these data in any supporting documentation. New guidelines for these managing these discussions are currently being drafted. The focus is to gather as much relevant information as possible at the time collections are acquired. Protocols such as those listed below are among the suggested practices:

- Identifying all known individuals contained in items in the collection
- Identifying all known authors and/or creators
- Identifying all known third party rights holders
- Identifying permissions and/or releases from third party rights holders
- Identifying documentation related to rights such as releases, contracts, etc.

To the best of its ability the Library is also aiming to secure rights as early as possible in the life of a digital object – or in some cases, the life of a physical object since this informs almost all future decisions. This is a change in practice that is being implemented slowly but successfully in the Research Libraries.

There are many ripple effects to these new practices. In the future the decision to acquire or not acquire a collection may rest of the status of digital rights. Furthermore, the collection development of the entire Library may shift because of these factors. Certainly, the role of the collection development staff will change.

Summary

The Library's role as a content creator, provider and distributor in the age of podcasts and webinars, wikis, blogs and Twitter is becoming critical. Access to new content areas is challenging especially when intellectual property rights are involved. However, reluctance by rights owners can be overcome by providing written transcriptions as memory aids to help make informed choices. Persistence is key. Often a donor's point of view can be swayed by gentle persuasion and expressed sensitivity to privacy issues.

Obtaining permissions and releases is essential for the success. If not already in hand, the process to obtain these allowances is slow and difficult and can impede the completion of the project.

The lion's share of the work with digital collections with rights that obtain to a party that is not the library is the proper recording and management of that rights information. All parties involved with these collections, including curators, directors, and counsel must share an understanding that rights information is not incidental, but essential to properly manage these collections.

Infrastructure – both technical and organizational - that supports the recording and management of detailed rights data should be implemented in libraries that wish to embark on similar projects. Obtaining these data as early on in the acquisition process as possible is optimal since researching these data is labor intensive and often precludes successful completion of these projects. Staff at all levels should be informed of these best practices and educated in the proper procedures for acquiring and documenting rights.

References

- ARALOC. (2010). *Cross platform DRM*. Retrived on May 2, 2010 from <http://araloc.com/>
- Berkeley Digital Library. (2007). *Copyright, intellectual property rights, and licensing issues*. Retrived on May 2, 2010 from <http://sunsite.berkeley.edu/Copyright/>

Discretix. (2010). *Multi-Scheme DRM client*. Retrived on May 2, 2010 from <http://www.discretix.com/DRM/index.html?source=adwordsCusGdrm-phrase888manufacturers100215&gclid=CKqHsMKX3KACFdlw5QodcyfZBQ>

New York Public Library Digital Gallery. (n.d.). Retrieved on May 2, 2010 from <http://digitalgallery.nypl.org>.

North Carolina ECHO. (n.d.). *Digitization guidelines*. Retrieved on May 2, 2010 from http://www.ncecho.org/dig/guide_1planning.shtml

Premis. (2010). *Preservation metadata maintenance activity*. Retrived on May 8, 2010 from <http://www.loc.gov/standards/premis/>

Stream Media. (2010). *Streaming Media Hosting is your expert for DRM - Digital Rights Management solutions*. Retrieved on May 2, 2010 from <http://www.streamingmediahosting.com/drm.htm?gclid=CMLolpWX3KACFWV75QodEAYXDA>

Washington State Library. (n.d.). *Digital best practice*. Retrieved on May 2, 2010 from <http://digitalwa.statelib.wa.gov/newsite/projectmgmt/vendors.htm>

Webpage Dodson, Howard African American Women Writers of the 19th Century. (1998). Retrieved on May 2, 2010 from http://digital.nypl.org/schomburg/writers_aa19/intro.html