Preserving Digital Public Television: Not Just an Archive, but a New Attitude to Preserve Public Broadcasting

Nan Rubin

> Public Television is responsible for the production, broadcast, and dissemination of programs which form the richest audiovisual source of cultural history in the United States.
> (Librarian of Congress, 1997)

Abstract

Television production has shifted rapidly from an analog process to one where virtually all programs are created and finalized as digital files. Such productions in public television are at great risk of being lost, because practices for long-term preservation of digital video are just now emerging, and because there is no mandate for preservation within the public broadcasting system. NDIIPP funded Preserving Digital Public Television, a partnership between WNET-TV in New York, WGBH-TV in Boston, PBS and New York University, to build a model preservation repository for digital video files and to examine the broader issues related to operating such a repository. In addition to designing the repository itself, the project became a lead advocate for adopting technical and metadata standards across the television field. The project also successfully challenged the public television system to recognize that preservation is necessary to keep digital productions alive. This resulted in public broadcasting allocating money for the first time to launch an initiative with a goal of properly managing its collective archival holdings.

From Analog to Digital: The Transformation of Television Production

In the first set of reports commissioned by the National Digital Information Infrastructure and Preservation Program (NDIIPP), the Library of

LIBRARY TRENDS, Vol. 57, No. 3, Winter 2009 ("The Library of Congress National Digital Information Infrastructure and Preservation Program," edited by Patricia Cruse and Beth Sandore), pp. 393–412

Congress identified the challenges of preserving digital television productions early in their development:

> By nature and necessity, public broadcasting is a hodgepodge of media types and formats . . . In whatever manifestations these objects previously existed, they become bits and bytes before they reach the public eye. That is an enormous amount of digital information to manage over time. As we move into the increasingly complex digital world, those charged with preserving our television heritage have the opportunity to develop and establish better coordinated and standardized preservation policies and practices to ensure what television programs and related assets survive. (Ide et al., 2002)

When this was written, it was not yet evident that television broadcast and production operations would be altered so profoundly or so rapidly by digital production and distribution technologies. In less than a decade, analog television has become totally obsolete by the availability of video recording and editing systems at prices within the means of most members of the public, the wholesale requirement of digital-only transmission, the ubiquity and immediacy of media on the Internet (in the form of podcasts, vodcasts, YouTube) and the mass deployment of handheld devices to view and listen to audio and video content.

In a relatively short period, television production and distribution has shifted from a linear, sequential, analog process based on physical media, to one that is almost entirely digital, which means easily programmable, random (nonlinear) access to content. Virtually all programs are now shot and edited in digital forms, and completed programs are finalized as digital files.

Distribution and transmission have been equally transformed. The Public Broadcasting Service (PBS) is replacing tape-based submissions for national program distribution with an operational system that will transfer digital files instead, using broadband networks able to handle large video files cost-effectively and with integrity. Likewise, nearly all local broadcast playback is now done in a tapeless environment, in which programs are stored as files then assembled and aired directly from a server.

The viewer environment has shifted as well. The Deficit Reduction Act of 2005 (Public Law 109-171, 109th Congress) requires all U.S. full-power television stations to turn off their analog transmitters by February 17, 2009 and begin broadcasting exclusively on digital channels. This means all analog over-the-air television signals will end, and Americans who want to watch television will need a digital television receiver of one kind or another. The all-digital television chain will be complete, from program producer at the front, to the viewer at home at the end of the line.

Not only have production and distribution methods been radically altered, the marketplace itself has changed. Statistics showing that a third of all Internet users ages eighteen to twenty-nine watch or download a

video online every day (Madden, 2007) challenge television's traditional concept of the audience. This has forced public broadcasting to examine nearly all its existing operating models, from over-the-air "appointment" program schedules and geographic market separation, to its traditional mix of support from government funding, corporate underwriting, private foundations, and viewer contributions.

As viewing has shifted away from television and onto the Internet, there has been particular scrutiny of how to expand the reach of public television content beyond the broadcast schedule. This has meant not merely putting finished programs online, but challenging stations and producers to use the Internet to highlight more video, share other program-related materials, and invite user-generated content.

In a culture that expects broadcast media to be available whenever it chooses, the notion of a video archive takes on new meaning: not as a barrier to accessing older content, but rather as a guardian protecting that content and keeping it vital.

Digitally produced programs in public television are at great risk of being lost, however, and not solely because of the rapid changes in technology that are rendering digital video recording and playback systems obsolete. The relatively short duration of program broadcast rights (often no more than five years) further complicate preservation, as rights associated with uses *after* broadcast are often absent. Combined with the great array of original source materials used in many programs and the complex rights status that may be attached to each show, the public television system faces a potentially tremendous financial burden to make programs viewable again.

Established practices that have served to archive and protect analog television programs on videotape cannot be used for the long-term preservation of digitally produced broadcast programs. Exponentially more elements are created in myriad formats throughout the lifecycle of a program, and, critical metadata are created in databases, spreadsheets, and on paper as well as many different types of electronic records. Leaving archiving to the end of the lifecycle opens the door to a host of threats including dissociation, interoperability and migration problems, and obsolescence, not to mention managing a volume of unused materials that far exceeds that of the analog world. A new approach that incorporates preservation practices into the entire digital production chain needs to be created.

Preserving Digital Public Television: A Significant Collaboration

As part of the Public Broadcasting Act of 1967, Congress authorized the Corporation for Public Broadcasting (CPB) "to establish and maintain, or contribute to, a library and archives of noncommercial educational and

cultural radio and television programs and related materials." However, CPB has never chosen to make that investment, and up to now, CPB funds have never been allocated for archival support.

Early agreements, such as those with the Library of Congress and the National Archives and Records Administration, have provided some preservation security for completed national programs aired through Public Broadcasting Service (PBS) and National Public Radio (NPR), but access is extremely limited. Moreover, local programs, which more closely reflect our daily lives, have not been included in these collections.

No mandate for system-wide preservation exists at any major public television institution, and there has been a notable lack of funds to invest in or allocate for preservation. Consequently, no single entity has the resources or expertise to manage this task by itself. Only a small number of stations and producers have been able to take on the responsibility and costs of preserving their own materials, so that within public broadcasting, very few formal archiving activities are in place. This leaves planning for digital preservation as an afterthought in the lifecycle of public television programs.

As one of the major producers of national programs in public television, Thirteen/WNET, New York's public television station, recognized this challenge and in partnership with WGBH in Boston, was awarded an National Digital Information Infrastructure and Preservation Program (NDIIPP) grant from the Library of Congress in 2004 to design a long-term preservation repository for "born-digital" public television programs.

These two television stations produce roughly 60 percent of the national prime time series that appear on public television, including signature public affairs and science programs such as *Frontline* and *NOVA,* which originate at WGBH, and leading cultural and historical materials, such as *American Masters* and *Great Performances*, which are produced at Thirteen/WNET. As "content creators," the stations control their own production units and thus have direct access to high-resolution "master" files of completed national programs.

In addition, WGBH was already recognized as a leader in identifying issues related to digital video preservation, and it brought a demonstrated history of developing and promoting digital asset management systems within public television.

WNET and WGBH maintain the only two station-based archives in the public television system, so each was also able to provide the project with professional staff experienced in preservation practices related to both digital and analog video materials.

The third partner was the Public Broadcasting Service (PBS), which operates the national network that distributes public television programs to more than three hundred stations for local broadcast. Because the majority of national programs pass through PBS before being aired lo-

cally, it is the principal de facto repository for this material, and the PBS warehouse holds more than 150,000 analog tapes of programs going back more than forty years.

These three institutions are directly engaged in both managing production workflows and holding the primary collections of public television programs seen by national audiences, and their interest in participating in NDIIPP was based on a shared recognition that public television had to begin taking steps to protect its rapidly growing collection of digital assets. Because they are primarily broadcasters, however, these institutions have few resources to operate digital libraries or develop preservation repositories.

New York University provided the expertise that was lacking in these areas. The NYU Digital Library team has extensive experience designing repository systems specifically for transferring and preserving large video files wrapped in rich metadata. The project further benefited from a relationship with NYU's Moving Image Archiving and Preservation Masters Degree Program, whose students have produced excellent research as part of the project, and whose graduates have become full-time project staff.

Project Goals: Build and Test a Model Repository

The intent of the project was to develop a small model repository that could be scaled up and leveraged for use by the larger public television system. Because public broadcasting as a whole has little exposure to preservation issues, it was also thought that the project could offer valuable resources to the system by drafting guidelines for content selection and appraisal, studying copyright impediments, and examining relevant financial and governance models. An unstated though larger and more important goal was to promote an understanding within public broadcasting that to exploit its programming well into the future, digital preservation should be a new priority worthy of investment.

In this context, Preserving Digital Public Television was designed as a series of discrete tasks to be tested in a lab-type environment. The initial set of activities specified:

- Designing a test repository for born-digital public television content.
- Developing a set of standards for metadata, file formats, wrappers, and production workflow practices.
- Drafting recommendations for appraisal policies for selecting public television content for inclusion in the repository.
- Examining issues of content accessibility and long-term operational sustainability.

Because the repository concept focused on capturing metadata during the production process, one early task was to examine production workflows. This would identify the points where key metadata was created at

various stages, and see if it was carried through the program lifecycle and ultimately used for preservation. Further activities centered on researching copyright issues involved in digital public television preservation, and capturing websites that are an extension of the program content offered over the air.

In planning for the NDIIPP project, the public television partners understood that identifying commonly used file formats and production protocols, determining appropriate metadata requirements, and adopting technical standards would have to be tested and collectively agreed upon. The project naively assumed that both commercial television networks and large collecting institutions like the Library of Congress, which was completing its new Packard Campus of the National Audio-Visual Conservation Center, were making progress on solving these same problems so that public television could benefit by contributing to work already underway.

The project quickly learned, however, that this was not the case. Other video producers including the networks, as well as the Library of Congress itself, were in fact struggling with the same technical issues, but no organization was making enough progress that public television could just follow along. Instead, the project found itself in the unanticipated position of leading the effort in the television industry to create a standard for video file wrappers, and adopting one of the first sets of metadata schema appropriate for long-term video preservation.

As for the intent to bring a new consciousness to the system, our project has had a broad impact among public broadcasters by sparking widespread support for launching local as well as national preservation efforts. The culmination of these activities is leading to the creation of the American Archive, a new national entity that has been proposed to provide preservation services for public radio and television program producers for the very first time.

Project Activities

Designing the Test Repository

The expertise of the NYU Digital Library team in building other archives with similar content was leveraged to design the repository architecture for this project. To test ingest and retrieval for the repository, WNET and WGBH selected a sample of individual high-resolution production master files from national program series that included hour-long episodes of *Nature* and *Frontline,* and episodes of *Religion & Ethics Newsweekly,* a half-hour each. Several episodes of the local half-hour WNET program *New York Voices* were also selected as a sample, representing non-national programming that did not go through PBS.

Before sending a program out to stations for broadcast, PBS processes

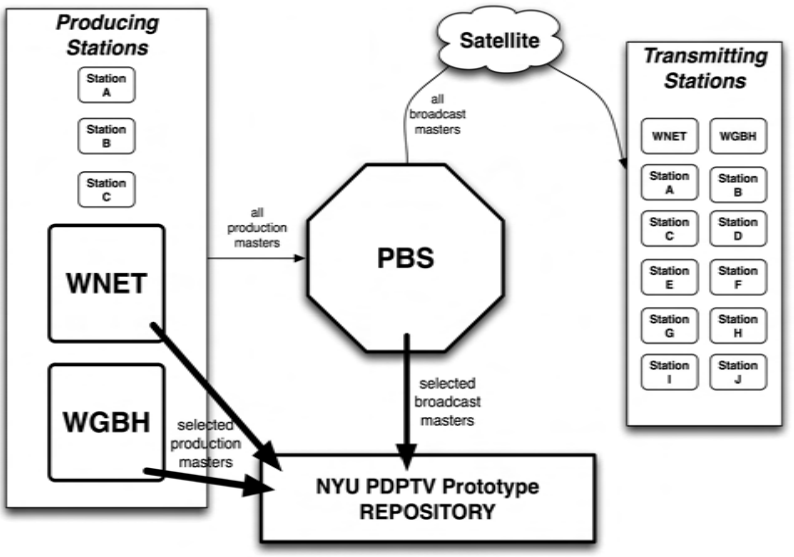

Figure 1. Repository Source Files

the file by adding elements such as underwriting announcements and compresses it for easier distribution by satellite or other means. To complement the production masters of the programs provided by WNET and WGBH, PBS provided the repository with the low-resolution distribution versions of the same programs (except *New York Voices*) (Figure 1).

This grouping allowed the project to test a mix of both high and low resolution program file formats created with varied workflows, accompanied by a wide range of metadata that was available but had to be collected separately from multiple sources. Altogether, more than thirty-five hours of programming was ingested into the repository, which included multiple versions of single programs stored in different formats.

The initial tests of the repository revealed that, despite the uniform requirements from PBS for submitting completed programs for national broadcast, the finished materials coming directly from the production units had to be ingested in several different flavors of formats and wrappers (MacCarn, 2007). These were not all equally easy to extract technical, metadata from or to read.

The project also found there was very little consistency in the way metadata for each program was recorded, as production units created certain elements of metadata, PBS generated others, all of which was created in different places for different purposes (i.e., production vs. distribution). This metadata was not collected consistently in a centralized place even within PBS, so collecting it to meet the needs of the repository had to be done on a program-by-program basis.

Determining Metadata Needs

Determining an appropriate set of metadata fields was a detailed and intensive task. Based on the assumption that the Preserving Digital Public Television repository should be OAIS-compliant, the project examined a broad range of standard metadata schema used by libraries and archives, as well as those emerging in the commercial television world. Project staff also reviewed PBCore (http://www.pbcore.org/), a metadata dictionary based on Dublin Core designed specifically for public radio and television program files.

One of the fundamental requirements of the repository was to aggregate content and metadata for a single program that came from disparate sources, both from the producing station (the high-resolution production master and database exports) and from PBS (lower-resolution broadcast master and more database exports).

Because the model repository does not input any new metadata, in order to create a useful Archival Information Package (AIP), sufficient information about each program file had to be packaged and sent along with the video as part of the Submission Information Package (SIP). The program SIPs therefore had to include appropriate metadata as well as the program files themselves (Figure 2).

To determine the metadata components required for the AIP, database exports from both program creators and PBS had to be analyzed, particularly the extensive descriptive and rights metadata created by PBS for broadcast scheduling.

Early on, the project partners chose PBCore to capture program metadata. PBCore has been in development for several years, but remains in the early stages of system-wide adoption. Even so, the repository designed its descriptive metadata requirements around PBCore, which has encour-

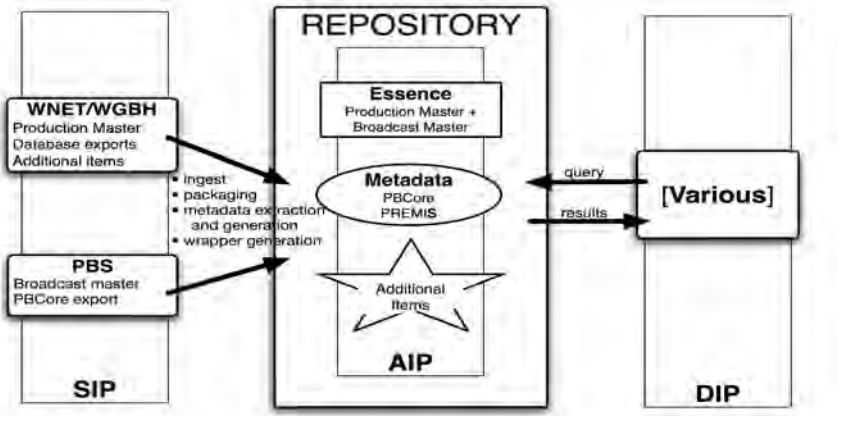

Figure 2. Repository Information Packages

aged other users to design and implement PBCore-compliant exports from their various databases. As a result of this effort, the most important source of metadata for national programming, PBS's Program Offer Data Service (PODS), can now be exported directly into PBCore, making information easy to ingest, package, and disseminate.

Incorporating rich technical metadata from the video files also proved to be a challenge. Because the program files were submitted to the repository in diverse formats (including a diversity of wrappers and encoding formats such as MXF, QuickTime, and various flavors of MPEG, and DVC Pro), multiple tools were required to play the videos and to extract technical metadata from the file headers such as bit rate, file size, and frame size.

Transforming the submitted metadata into a standard format was a clear necessity. The solution was to create a schema that encapsulated the necessary descriptive and technical metadata from a variety of different data dictionary standards, while maintaining information unique to public television programming.

To accomplish this, the repository has developed a structure to capture all necessary metadata using not only PBCore, but also PREMIS and METSRights. Appropriate fields from these standards, along with virtual links to the program files themselves, are all contained within a METS wrapper (Figure 3).

The problems that were encountered by testing these various file formats, combined with the time-consuming efforts needed to collect metadata, demonstrated the high priority for setting uniform standards as a requirement for the success of any future repository operation. Without

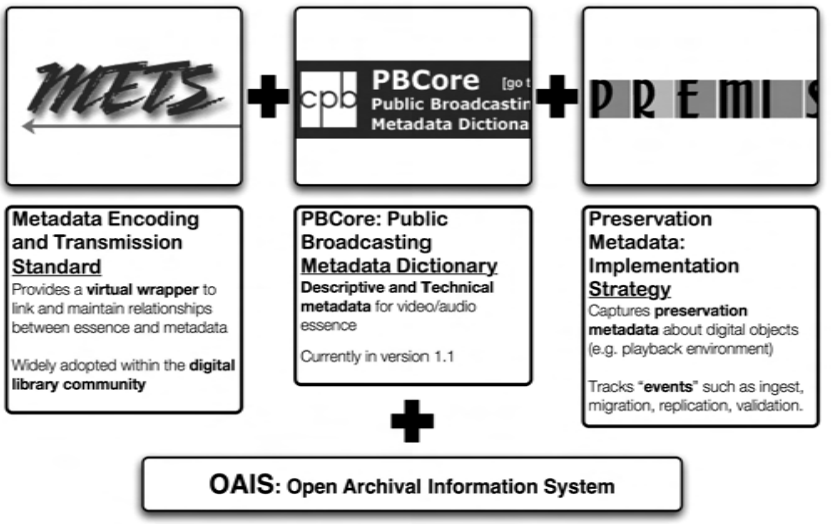

Figure 3. Metadata Structure

a set of accepted standards to apply to the creation of the program file formats and descriptive metadata, automating the functions for extracting and managing the metadata of large collections will simply not be feasible.

Video File Formats and Wrappers

The creation of a relatively stable standardized video file wrapper is needed to facilitate the successful exchange and long-term preservation of digital files, particularly to support future file migration. A number of such so-called video wrapper "standards" exist, but despite vendor claims, the files do not all actually interoperate with different equipment.

This problem was identified early as one that the project wanted to address, because there was no single acceptable wrapper when the project was initiated. To encourage a shared approach to solving the problem, the project convened a group of more than twenty distinguished technologists, digital collections designers and project managers. At the "Wrapper Round Table," the public television partners were surprised to learn that the lack of consistent video format and wrapper standards was also a major problem shared by the commercial broadcast networks (Fleischhauer, 2006).

To be successful, any initiative to create standards for broadcast files must dovetail with the needs of the commercial broadcasting industry because public television on its own does not carry enough economic clout to influence hardware vendors to adopt and support the desired standards.

This led members of the NDIIPP team to a series of exploratory meetings with such diverse groups as Turner Broadcasting, the Discovery Channel, and the Department of Defense, and eventually resulted in the Advanced Media Workflow Association (formerly the Advanced Authoring Format Association, which represents system vendors) launching the AS-03 wrapper project specifically to ensure "use of MXF as the file wrapper to facilitate a vendor-neutral format to support interoperability" (2008, July 18).

Analyzing Program Production Workflows

With the understanding that preservation and other valuable metadata must be captured early in the process and carried through the entire program lifecycle, another aspect of the project was to examine production workflows and identify the points where key metadata are created so that preservation practices could be integrated throughout.

The different workflows associated with the test programs were analyzed to identify how this might be done, and although they could not be altered, the workflows in place at PBS for accepting and preparing program files for distribution were also examined.

Production workflows are complex and varied, depending on the type of program and its style. For example, *Religion & Ethics Newsweekly* is a current affairs magazine comprised of several different segments packaged

together to create one broadcast episode each week. It both produces its own segments and acquires materials from other sources, and like many news programs, it must be edited and assembled quickly, with programs often being completed barely by deadline. By contrast, *Nature* programs are often shot and produced by just one entity, with a deadline of a year or longer.

Each production unit has its own practices and keeps records in its own way, and to get an idea of how many different program production units exist, just look at your weekly public television schedule. There are as many units as programs. The volumes of information accumulated during a production are a mix of paper and electronic records based on legacy workflows learned or inherited by the producer and then added to by PBS, with no uniform databases and no standardized recordkeeping shared between or across any productions.

There is no "one size fits all" solution for incorporating preservation practices into such diverse workflows, but we initially thought that a suitable approach might be adapted to test on one or two programs. This idea proved both elusive and intractable.

Apart from the inefficiencies of production workflows that largely replicate analog procedures in a digital environment, it was simply not possible to propose to any production units of our selected test programs that they alter their workflows to accommodate the needs of long-term preservation. By and large, individuals involved along the production path are not inclined to see the value of adding additional work to their already demanding tasks, especially if they are working against a deadline.

More importantly, the tools necessary to make these changes were not in place. Without a mechanism that makes it easy to ingest, store, browse, and view digital video, creators do not have enough incentive to change their practices. But currently, neither WNET nor WGBH has an appropriate digital asset management system (DAM) to manage the digital production workflow. Until such a system for storing program files, content elements, and metadata during the production process is implemented, workflows cannot be significantly altered.

Content Selection and Appraisal

On a practical level, examining issues related to selecting public broadcasting content and drafting recommendations proved to be one of the easier aspects of the project to complete. Through a series of focus groups, teachers, educators, journalists, documentarians, academics, and others who use television in their work provided their input.

The project partners also conducted a literature search, and collected operating policies from a number of film and video archives. This research resulted in the publication of *Recommended Appraisal Guidelines for Selecting Born-Digital Master Programs for Preservation and Deposit with the Library of Congress* (Ide and Weisse, 2006). This report outlines existing and emerg-

ing practices in moving image appraisal that are appropriate for use with national public television programming.

The project team determined that local stations wanted additional guidance in appraising their own local program productions. *Appraisal and Selection Guidelines for Public Television for Public Television Volume II: Criteria and Recommendations for Local Stations, Producers and National Productions* (Preserving Digital Public Television, 2008) focuses on specific criteria to appraise not solely national materials, but also elements and program versions of local productions.

Knowing that the Library of Congress already had a vested interest in collecting the "best copy" of programs distributed nationally by PBS "in complete and unedited broadcast format . . ." (PBS-Library of Congress Donor Agreement, 1993) this also shaped the selection recommendations, so that recent programs like *The War* and *African American Lives* will be able to join the ranks of long-running series like *NOVA*, *Live from Lincoln Center,* and *P.O.V.,* along with *An American Family, Julia Child, The Adams Chronicles, Great American Dream Machine,* and other older classics as important documentation of the U.S. experience.

Managing Program Rights

Along with issues of storage and technical architecture, the most common challenge in preserving digital video concerns intellectual property and related rights. Program productions are a complex amalgamation of original moving images, stills, unique art, scripted narrative, dialog, recorded sound, music, performance, acquired footage, and many other elements. Each part can potentially carry a long list of underlying rights—ownership and creator rights, restrictions and authorships, union contracts, distribution agreements, and other specific conditions for use. Typically, a producer obtains the right to use a clip of historic footage or music in a television production for a period of only five to ten years. In addition, many public television funding models include co-ownership or shared copyright of programs to relieve the financial burdens of production. One or more entities might therefore own the copyright on a single finished program, while many others hold the mass of underlying rights.

Distribution rights allow public television programs to be broadcast a maximum number of times on noncommercial television during a finite window, for example, six showings over the air in five years. Generally, rights for a certain level of home video sales and educational use in the K-12 classroom are also part of the distribution rights package.

When these rights expire, the system no longer maintains much interest in the content unless there is an exceptional demand from viewers. Then the rights to broadcast the program again have to be "re-upped" or renewed for air by getting permission from the copyright holder as well as from all the underlying rights holders.

Reuse of the program through any new distribution method such as via the Internet, or for an additional length of time, is often discouraged because attaining the rights to do so can be very difficult. Identifying and locating all the underlying rights holders can entail a great deal of research, and paying for permissions to clear the program for reuse could easily be quite substantial.

For many productions, particularly those that use archival or stock photos or clips, music, etc., this can be very costly both in terms of research to find all the rights holders, and any direct costs that may be required. For example, the important civil rights documentary series *Eyes on the Prize* could not be broadcast or otherwise distributed for a decade because the producers could not afford the cost of renewing all the underlying rights (Dean, 2005; Bernard, 2005).

Specific authorization to preserve national public television programs is largely absent, and seeking permission to put programs into a digital repository for preservation and access is too emergent an idea to have much precedent. At the same time, prompted by public expectations that television programs should be readily available online, many rights holders are trying to create new restrictions to maintain control of their creations out of fear of losing the opportunity to exploit their own materials.

Under these conditions, the model repository would have to stay largely "dark" unless it has explicit permission from the copyright holder(s) to allow programs to be used for anything other than archival research.

The Preserving Digital Public Television project team decided that a discussion paper relating to access and rights is critical to inform plans for a functional repository in the near future. The paper will include recommendations for drafting new model language for contracts and distribution agreements to authorize appropriate perpetual rights for long-term preservation and noncommercial access to the content.

Planning for Sustainability

A major functional issue for any digital repository is how it will be economically sustained, what expenses will be required for staffing, administration, services, and maintenance along with technical facilities, storage, interconnection, migration, and the like. As existing digital repositories mature, the wide ranges of these costs are beginning to be documented and analyzed (see Additional Resources). The National Science Foundation, in partnership with the Library of Congress and other large research institutions, is also studying cost models and issues of sustainability.[1]

Because the argument for program preservation is not universally supported in public broadcasting, a key challenge in discussing sustainability is to rely not solely on numbers, but to offer a compelling rationale for making this investment in the first place.

Given the relatively few sources of income traditionally available to

public broadcasting, the economics of setting up and operating a preservation repository might seem overwhelming and unaffordable. Instead, the public television community must be reassured by seeing it as both feasible and viable. The case for sustaining a repository must be made without putting it in competition with the immediate demands of station operations and program production, and without being perceived as overly intimidating.

Our contribution to this effort has been to focus on solving the particular problems of maintaining very large digital video files, and to keep the projected scale manageable by describing a service that would initially provide only basic functions but system-wide benefits. The project is also closely monitoring the new research on this topic being published with growing frequency, and will incorporate the most relevant findings in our own projections.

Preserving Websites Because Programs Aren't Enough

Websites are where the public turns to for up-to-date information, resources and background materials, and they have become required components for every public television production, both local and national. Websites provide expanded information about any given program, and increasingly they reflect large amounts of original video and other cultural expression. Public broadcast stations themselves have active websites, which provide media and other content that ties them directly to their communities.

Increasingly, websites contain far more video than can be contained in any single broadcast, and they are now considered an integral aspect of the broadcast record. As such, preserving them had to be considered along with saving the programs themselves. With the Internet Archive, at the end of 2007 the project began saving a large collection of websites related to the public television system, including local PBS stations across the country and individual program productions and series. Nearly four hundred Web addresses are being captured on a monthly basis (PTV Digital Archive, n.d.).

Public television's Web presence will be transformed during 2008 as the system adopts new Web tools that will greatly increase the amount of video offered online. In response, many stations and producers are redesigning and reorganizing their public websites. Our Web crawling activities will be documenting the evolution of these websites and the impact the changes are having on improving the interaction of public television with its viewers. As such, they will provide an important additional window into the larger social landscape reflected in the content distributed over the air.

Impact of Non-Broadcast Digital Distribution

In designing the repository, the project began confronting important dynamics with impact well beyond the initial scope of work. One of the most significant dynamics has been the transformation of television program-

ming distribution by digital technology, in tandem with viewers' near-universal access to the Internet. These shifts have resulted in rapid migration away from over-the-air "appointment viewing," to on-demand viewing of programs saved on DVDs, downloaded, or streamed online.

There has also been an explosion of demand for older broadcast content to be findable online, and an expanding library of such material is becoming steadily available.

When coupled with a constantly changing array of viewing devices, these phenomena have created a fundamentally altered video environment that requires programming to be available as digital files viewable on the very smallest iPod screen, to wall-size flat panels.

Such factors have forced public television to face the question of how to release both current and older program materials into the ocean of online video offerings. Producers began to understand that to be successful amid such extremely fluid technology, they had to come to terms with the need for adopting standards for file formats and file storage; saving appropriate metadata; resolving certain rights issues; and acquiring effective digital content management tools.

Stations began to look for best practices and cost-effective solutions to solve these online distribution problems. They found that, although our NDIIPP project had approached these questions specifically through the lens of long-term preservation, most of the problems relating to creating a stable preservation environment overlapped significantly with their own efforts to get their video online.

Preserving Digital Public Television thus became highly relevant to television stations and producers across the system, and in less than two years the whole concept of digital preservation moved from being a marginal concept directly into being a key factor in the core public television debate on how best to make content available to reach more viewers.

Leading Within the Public Broadcasting System

The steady rise of nonbroadcast digital distribution led to a genuine interest in the NDIIPP project. This was not only because of the expertise the project was developing in the technology realm, but also because the project brought with it a wide array of useful resources and contacts from far outside the public broadcasting sphere. Also, through the project, public television was becoming a valuable preservation partner to such institutions as the Library of Congress, the National Science Foundation, the National Archives and Records Administration, and the Academy of Motion Picture Arts and Sciences. The project was helping position the system to participate in significant new funding opportunities.

From the beginning, the project intended to build support by engaging the public broadcasting community. In the early stages, this meant primarily going to public television's annual technical gathering and meeting with station engineers and technologists informally to solicit in-

put. As the system shifted more firmly into an all-digital environment, we became more active promoting the value of digital preservation and organizing broader station participation in industry gatherings.

By 2007, public broadcasters began seeking us out to present at conferences and symposia, and to provide guidance on how to approach preservation planning. When the article "Everything Old Can Be New Again" (Rubin, 2007) was published, we became the most visible effort promoting preservation, introducing stations and producers to acceptable preservation practices and advancing system-wide collaborations to support PBCore and other technical and metadata standards.

For example, because of the project's high profile, WHYY Radio in Philadelphia asked the project for advice on preparing its library of twenty-five years of audiotapes of the daily program *Fresh Air with Terry Gross* for digitizing and cataloging. The station was especially interested in adopting appropriate public broadcasting standards such as PBCore, and did not want to proceed without aligning its efforts with other similar projects in public radio and television.

Major project efforts have also encouraged local broadcasters to create preservation partnerships with cultural heritage organizations in their own regions that share needs for creating and managing digital collections. Inspired by this message, the assistant general manager for Content at NET (Nebraska's state-wide distance learning and public broadcasting network) organized his own regional meeting, bringing together more than thirty participants from NET's Broadcast and Web services, the University Library School, the Nebraska Regional Humanities Center, the Nebraska State Historical Society, and the Historical State Archivist, to explore interest in setting up some form of digital preservation collaboration.

The American Archive: A New National Initiative

These examples illustrate the success of Preserving Digital Public Television in inserting the framework for preserving digital video into the immediate environment of stations operations and program production. But the most dramatic response has come from the Corporation for Public Broadcasting (CPB) itself, which in 2007 initiated a new project to create what is being called "The American Archive." The concept of the American Archive, as outlined by Congressman Ed Markey, is "to harness the power of digital technology . . . to preserve public broadcasting's audio, film, and video history, and to make it available to the American people" (Speech to the Association of Public Television Stations, February 2007).

For the first time, CPB is taking steps to invest funding in preservation activities. The American Archive is being planned as a totally new entity, independent of existing public broadcasting institutions like National Public Radio or PBS, in part as a response to growing momentum within public broadcasting from projects like ours.

It will take several years to shape the American Archive, outline how it will operate, design a governing structure, and build sustainable funding streams. Nonetheless, public television stations across the country have endorsed the project and are now organizing their own resources to contribute to the collective effort.

Maintaining the Momentum

Since this project began, public broadcasting has shed its analog systems and moved completely into a digital universe. Along with the emergence of cost-effective digital technologies for content management and storage, the changes have involved more than swapping tape machines for computer servers, but also adopting totally new approaches to editing, sharing, and maintaining program content.

By continuing to have high visibility, Preserving Digital Public Television has been able to impress on the public television system the message that digital preservation practices cannot be an afterthought, but are important components to be integrated into the entire program production lifecycle. The project clearly identified the most basic problems that had to be solved, with success in meeting some better than others, and tied them directly into the broader operating concerns that the system is facing.

On a more basic level, the project successfully challenged public television to discuss the need for preservation in order to keep digital productions alive at all. This had the impact of transforming an attitude of indifference across the system to one that recognized the importance of properly managing our collective digital archival holdings.

In this, Preserving Digital Public Television has played a major role preparing the groundwork for long-term collaborations that have the potential for building widespread support for digital preservation activities. But we cannot go much further without broader system involvement.

The public broadcasting community is ready to express its desire to keep our television content vibrant and useful. The technical conditions necessary to operate preservation repository will soon be solved, and the behaviors needed to add preservation-relevant metadata into program production and distribution workflows will eventually be adopted. What remains is that public broadcasting organizations continue building commitment, so that once preservation projects like ours are underway, they will be sustained and nurtured over time, not solely to serve the system, but because they benefit the American public as a whole.

Acknowledgements

Special thanks to the *Preserving Digital Public Television* NDIIPP team for their assistance in preparing this article, especially Dr. Howard Besser, professor and director Moving Image Archive & Preservation Program, NYU; Kara van Malssen, senior research scholar, NYU; Yvonne Ng, junior research scholar, NYU; Karen Cariani, director, WGBH Media Library

and Archives; and Bea Morse, director, Business and Communications, PBS Interconnection Replacement Office.

Note

1. Funded by the National Science Foundation and the Andrew W. Mellon Foundation, in partnership with the Library of Congress, the Joint Information Systems Committee of the United Kingdom, the Council on Library and Information Resources, and the National Archives and Records Administration, the Blue Ribbon Task Force on Sustainable Digital Preservation and Access was created in late 2007. During the next two years, the BRTF-SDPA will explore sustainability and deliver specific, economically viable recommendations of use to a broad audience, from individuals, to institutions and corporations to cultural heritage centers.

Additional Resources

Preserving Digital Public Television

PTV Digital Archive. Retrieved January 15, 2009, from http://www.Ptvdigitalarchive.org

Moving Image Selection and Appraisal

Besser, H. (2001, Fall). Digital preservation of moving image material? *The Moving Image, 1*(2), 39–56.

Connors, T. (2000, Spring/Summer). Appraising public television programs: Toward an interpretive and comparative evaluation model. *American Archivist, 63*, 152–174.

Fleischhauer, C. (2003, Fall). Looking at preservation from the digital library perspective. *The Moving Image, 3*(2), 96–100.

Gray, F., & Sheppard, E. (2004, Fall). Moving history: Promoting moving image archive collections in an emerging digital age. *The Moving Image, 4*(2), 110–118.

Ide, M., & Weisse, L. (2003, Spring). Developing preservation appraisal criteria for a Public Broadcasting Station. *The Moving Image, 3*(1), 146–157.

Kula, S. (1983). Archival appraisal of moving images. Introduction to selected guidelines for the management of records and archives: A RAMP study with guidelines. PGI-83/WS/18. Paris: United Nations Educational, Scientific, and Cultural Organization. Retrieved January 15,2009, from http://unesdoc.unesco.org/images/0005/000576/057669e.pdf

Kula, S. (1997, March). Archival appraisal of moving images. In H. Harrison (Ed.), *Audiovisual archives: A practical reader* (Section IV, Selection and Appraisal 1.3). Retrieved January 15, 2009, from http://unesdoc.unesco.org/images/0010/001096/109612eo.pdf

Lewis, A. (n.d.). Archives notes from the public television archives. In *An occasional newsletter of the Public Television Archives of PBS* (p. 2). 2nd newsletter (Washington, DC: PBS).

Library of Congress. (1997). Television and video preservation 1997: A report on the current state of American television and video preservation. Washington, DC: Library of Congress.

Preservation File Formats for Video and Associated Metadata

Berkeley Multimedia Research Center. (n.d.). MPEG-2 FAQ. Retrieved January 15, 2009, from http://bmrc.berkeley.edu/frame/research/mpeg/mpeg2faq.html

Brown, A. (2003). *Digital preservation guidance note 1: Selecting file formats for long-term preservation*. London: The National Archives. Retrieved January 15, 2009, from http://www.nationalarchives.gov.uk/preservation/advice/pdf/selecting_file_formats.pdf

Dance Heritage Coalition and Media Matters, LLC. (2004). *Digital video preservation reformatting project*. Washington, DC: Dance Heritage Coalition. Retrieved January 15, 2009, from http://www.danceheritage.org/preservation/DigitalVideoPreservation1.pdf

Digital Library Federation. (2005). METS schema & documentation. Washington, DC: Library of Congress. Retrieved January 15, 2009, from http://www.loc.gov/standards/mets/

Factors for Evaluating Preservation File Formats for Digital Video. (n.d.). Retrieved January 15, 2009, from http://www.ptvdigitalarchive.org/docs/Technical%20Issues/Preservation%20File%20Formats%20-%20chart.pdf

Lawrence, G. W., Kehoe, W. R., Rieger, O. Y., Walters, W. H., & Kenney, A. R. (2000, June).

Risk management of digital information: A file format investigation. Washington, DC: Council on Library and Information Resources. Retrieved January 15, 2009, from http://www.clir.org/pubs/reports/pub93/pub93.pdf

Material exchange format (MXF). (n.d.). Retrieved January 15, 2009, from http://www.digitalpreservation.gov/formats/fdd/fdd000013.shtml

Library of Congress. (n.d.). MPG-2 video encoding (H.262). Retrieved January 15, 2009, from http://www.digitalpreservation.gov/formats/fdd/fdd000028.shtml

Library of Congress. (n.d.). MPEG-4 file format, version 2. Retrieved January 15, 2009, from http://www.digitalpreservation.gov/formats/fdd/fdd000155.shtml

Library of Congress. (n.d.). Sustainability of digital. Retrieved January 15, 2009, from http://www.digitalpreservation.gov/formats/index.shtml

Shepard, T., & MacCarn, D. (2001). *The universal preservation format: A recommended practice for archiving media and electronic records.* Boston: WGBH Educational Foundation.

Digital Sustainability

Baker, M., et al. (2006, April 18–21). A fresh look at the reliability of long-term digital storage. EuroSys'06, Leuven, Belgium. Retrieved January 15, 2009, from http://lockss.org/locksswiki/files/Eurosys2006.pdf

Hoorens, S., & Rothenberg, J. (2008). The uncertain future of saving the past. Rand Corporation.

Lavoie, B. (2008). The fifth blackbird: Some thoughts on economically sustainable digital preservation. *D-Lib, 14*(3/4). Retrieved January 15, 2009, from http://eprints.ucl.ac.uk/4914/

National Archives and Records Administration. Frequently asked questions (FAQs) about selecting sustainable formats for electronic records. Retrieved January 15, 2009, from http://www.archives.gov/records-mgmt/initiatives/sustainable-faq.html

Palm, J. (1996, May). The digital black hole. Preserving digital information report of the task force on archiving of digital information. Retrieved January 15, 2009, from http://www.tape-online.net/docs/Palm_Black_Hole.pdf

The Science and Technology Council of the Academy of Motion Picture Arts and Sciences. (2007). The digital dilemma: Strategic issues in archiving and accessing digital motion picture materials. Academy of Motion Picture Arts and Sciences, 2007.

Smith, A. (2003, May). Issues in sustainability: Creating value for online users. *First Monday, 8*(5).

Society of American Archivists. (2007). Constructing sustainability: Real-world implementations of preservation standards for born-digital design documentation. Chicago: Society of American Archivists.

Watson, J. (2005).LIFE project research review: Mapping the landscape, riding a life cycle. *Literature review.* London, UK. Retrieved March 20, 2009, from http://eprints.ucl.ac.uk/1856/1/review.pdf

Watson, J. (2006). LIFE: Bibliography. Retrieved March 20, 2009, from University College London, Eprints Publications: http://eprints.ucl.ac.uk/1857/1/LIFE_biblio.pdf

Wheatley, P., Ayris, P., Davies, R., Mcleod, R., Shenton, H. (2007). The LIFE Model V1.1. Discussion Paper. Retrieved March 20, 2009, from LIFE Project, University College London, Eprints Publications: http://eprints.ucl.ac.uk/4831/1/4831.pdf

References

Advanced Media Workflow Association. (2008, May 8). *Technical work proposal: Application specification for program delivery* (AS-03). Retrieved July 21, 2008 from Advanced Media Workflow Association, AMWA Engineering & Users Group Meeting Web site: http://www.aafassociation.org/passed_events/2008-05_EngMtg-ArlingtonVA/AMWA-AS-03-Delivery-projectproposal-20080508.pdf

Advanced Media Workflow Association. (2008, July 18). AMWA engineering project summary. Retrieved July 21, 2008, from http://www.aafassociation.org/html/projects.html

Bernard, S. (2005). Eyes on the Rights; The Rising Cost of Putting History on Screen. *International Documentary,* June, 26-32. Retrieved July 21, 2008 from http://www.harborproductions.com/print_files/IDAeyes6-05.pdf

Corporation for Public Broadcasting (2005). *PBCore v 1.1.* Retrieved July 21, 2008, from http://www.pbcore.org/

Dean, K. (2005). Cash Rescues Eyes on the Prize. *Wired Magazine,* August 30. Retrieved July 21, 2008 from http://www.wired.com/entertainment/music/news/2005/08/68664

Fleischhauer, C. (with Pearson, G. & Rubin, N.) (2006). *Wrapper Roundtable: March 9, 2006.* Retrieved July 21, 2008, from http://www.ptvdigitalarchive.org/docs/Wrapper_Roundtable_Notes.pdf

Ide, M., MacCarn, D., Shepard, T., & Weisse, L. (2002). Understanding the Preservation Challenge of Digital Television. In *Building a National Strategy for Preservation: Issues in Digital Media Archiving* (pp.67-79). Washington, DC: Council on Library and Information Resources and the Library of Congress.

Ide, M., & Weisse, L. (May 25, 2006). *Recommended Appraisal Guidelines for Selecting Born-Digital Master Programs for Preservation and Deposit with the Library of Congress.* Retrieved July 21, 2008, from http://www.ptvdigitalarchive.org/docs/Selection/Recommended%20Appraisal%20Guidelines.pdf

Librarian of Congress (1997). *Television and Video Preservation 1997: A Study of the Current State of American Television and Video Preservation.* Washington, DC: Library of Congress.

MacCarn, D. (with Rubin, N.) (2007). *Survey of digital formatting practices in public television program production.* Retrieved July 21, 2008, from http://www.ptvdigitalarchive.org/wp-content/uploads/ 2007/09/report-on-file-formats-and-packages-fin.pdf

Madden, M. (July 25, 2007). *Online video.* Retrieved July 21, 2008, from Pew Internet & American Life Project Web site: http://www.pewinternet.org/pdfs/PIP_Online_Video_2007.pdf

PTV Digital Archive. (n.d.). Web crawl. Retrieved July 21, 2008, from http://ptvdigitalarchive.org/our-work/web-crawl/

Public Broadcasting Act of 1967, 47 U.S.C. 396.

Preserving Digital Public Television (2008). *Appraisal and selection guidelines for public television for public television volume II: Criteria and recommendations for local stations, producers and national productions.* Manuscript in preparation.

Rubin, N. (2007, May 14). Everything old can be new again. *Current.org.* Retrieved July 22, 2008, from http://www.current.org/dtv/dtv0708preservation.shtml

---

Nan Rubin has more than thirty years experience managing technology and facilities projects in public broadcasting. She has worked at public television station Thirteen/WNET-TV in New York City on such major efforts as restoring the station's analog and digital broadcast signals after the destruction of the World Trade Center, and she planned and coordinated the creation of Thirteen's Tape Archives to preserve both analog and digital program assets. Since 2003, Ms. Rubin has been project director of Preserving Digital Public Television, funded by the Library of Congress's National Digital information and Infrastructure Preservation Program (NDIIPP), where she oversees a team of twenty based at Thirteen, WGBH in Boston, the Public Broadcasting Service, and New York University. Her article "Everything Old is New Again" on the importance of preservation planning has been distributed widely, and she has been an invited speaker at the Digital Library Federation, the National Library of Medicine, the PBS Technology Conference, the Blue Ribbon Task Force on Sustainable Digital Preservation and Access, and the International Federation of Television Archives.