

Sound Directions

Digital Preservation and Access for Global Audio Heritage

Statement for the National Recording Preservation Board
January 31, 2007

Submitted by:

Mike Casey
Associate Director for Recording Services
Archives of Traditional Music
Indiana University

Section I of this document presents basic information about the *Sound Directions* project, particularly at Indiana University, to provide context for the discussion of issues and concerns that follows in Section II.

I. Introduction

Sound archives have reached a critical point in their history marked by the simultaneous rapid deterioration of unique original materials, the development of expensive and powerful new digital technologies, and the consequent decline of analog formats and media. It is clear to most sound archivists that our old analog-based preservation methods are no longer viable and that new strategies must be developed in the digital domain. Motivated by these concerns, in February 2005 the Indiana University Archives of Traditional Music (ATM) and the Archive of World Music (AWM) at Harvard University began Phase I of *Sound Directions: Digital Preservation and Access for Global Audio Heritage*—a joint technical archiving project with funding from the NEH Preservation and Access Research and Development grant program. The goals of Phase I of *Sound Directions* were to a) create best practices and test emerging standards for digital preservation; b) establish, at each university, programs for digital audio preservation that will enable us to continue this work into the future; and c) preserve critically endangered, highly valuable, unique field recordings of extraordinary national interest.

Sound Directions focuses on field recordings—carriers of unique, irreplaceable and historically significant cultural heritage. As caretakers of these collections we must solve the problem of preserving audio resources accurately, reliably, and for the very long term; at the same time we must make our resources readily accessible to those who most need them. These issues have been the subject of work, discussion and study at a number of national agencies and institutional archives, including the Council on Library and Information Resources, the American Folklife Center, the Library of Congress Audio-Visual Prototyping Project, and, as stated above, *Sound Directions* Phase I. Most of us are now approaching audio digitization in similar, deliberately cooperative ways. One key contribution of Phase I of *Sound Directions* has been the creation of interoperable digital audio preservation packages.

The *Sound Directions* project at Indiana University is a collaboration between the Archives of Traditional Music and the Digital Library Program (DLP). The ATM is one of the largest university-based ethnographic sound archives in the United States. Its holdings cover a wide

range of cultural and geographical areas, and include commercial and field recordings of vocal and instrumental music, folktales, interviews, and oral history, as well as videotapes, photographs, and manuscripts. For over fifty years, the ATM has been a recognized leader in the sound archiving community, developing in step with technological and theoretical advances in ethnographic research and recorded sound. The ATM holds some 70,000 original recordings, many of which were recorded in the field from the late 1890's to the present and are in various stages of deterioration. Some formats held by the ATM, such as lacquer discs, are degrading rapidly with loss of content already a reality. Two examples will illustrate the content that we are losing:

- Lacquer disc 16-601 has cracks in the coating from delamination on some of its bands, rendering much of it unplayable. This disc contains performances by a stringband recorded in 1947 by radio station WOWO in Fort Wayne, IN. Many of the discs from this particular collection are delaminating. Tape copies of parts of the collection were made before 1962 on unstable acetate-based tape. We have no way of knowing if the contents of disc 16-601 are included on the tapes. The opportunity to undertake preservation transfer of this material to high resolution digital files is lost forever
- Two open reel tapes from renowned folklorist Henry Glassie's Collection—OT 11102 and 11103—are currently unplayable due to problems with extreme squealing and stiction. These tapes are *not* suffering from Sticky Shed Syndrome, which has a temporary fix available. The tapes were recorded in 1972 and contain political songs, stories, and instrumental dance music from Northern Ireland. While we hope eventually to find or develop a procedure to restore this material, this content may currently be considered lost.

With deterioration continuing, and accelerating in some cases, strategic selection for preservation becomes vitally important. Selection at the ATM is undertaken using a software tool named FACET, developed in part by the *Sound Directions* project, which assesses the condition of field collections and the level of risk they carry.¹ Data from this tool is combined with both an analysis of research value and the experience of staff to produce priorities for preservation work. Although the preservation system developed at the ATM is designed to be sustainable, there is currently only one audio engineer, supported by grant funding, who engages in preservation transfer work for this very large and endangered recorded sound archive. Obviously, this program is not yet self-sustaining.

ATM's partner in this project, the IU Digital Library Program, is dedicated to the selection, production, and maintenance of a wide range of high quality networked resources for scholars and students at Indiana University and elsewhere, and supports digital library infrastructure for the university. The DLP is a collaborative effort of the Indiana University Libraries, the Office of the Vice President for Information Technology, and IU's research faculty with leadership from the School of Library and Information Science and the School of Informatics. DLP staff provides expertise in planning, creating, and maintaining digital projects. DLP's Variations2 digital music library project received a \$3 million grant from the National Science Foundation to create an integrated digital library that presents users with access to sound recordings, musical scores, and video in a variety of formats. A recent \$768,000 grant from the Institute of Museum and Library Services for the Variations3 project is extending this digital music library to institutions beyond IU. The IU Digital Library Program is also in the process of designing and

developing a new technical infrastructure for the storage and delivery of digital library collections, centered around a preservation and access repository implemented using the Fedora open source digital repository platform from Cornell University and the University of Virginia. By late 2007, the DLP will implement a bridge between Fedora and the university's Massive Data Storage Service (MDSS, described below) so that the repository may make use of this service for replicated storage of very large files.

The Massive Data Storage System is a distributed storage service offered by Indiana University's University Information Technology Services. This system is based on a consortium-developed software product known as High Performance Storage System and consists of nearly 1.6 petabytes of disk and automated tape storage. Through use of the State of Indiana's I-Light optical fiber network, MDSS supports automatic mirroring of data between servers on IU's Bloomington and Indianapolis campuses for added disaster- and fault-tolerance.

Key personnel for the *Sound Directions* project at Indiana University include Daniel Reed, Director of the ATM, who is co-Principal Investigator. Mike Casey, Associate Director for Recording Services at the ATM, serves as project manager for the Indiana portion of the project and adds his perspective as co-chair of the ARSC Technical Committee to this document. At the Digital Library Program, Jenn Riley, Metadata Librarian, and Jon Dunn, Associate Director for Technology, work on the preservation repository side of the project.

It is within this context that the Indiana University portion of the *Sound Directions* project offers the following comments on the state of audio preservation in the United States, outlining some of the critical issues that we feel must be addressed for this field to successfully move forward.

II. Comments and Discussion of Issues

A. Preservation Systems

Preservation operations may be conceptualized as systems, with each part ideally contributing to an efficient workflow that maximizes throughput while resulting in high-quality output of preservation-worthy digital objects. An effective preservation system in the analog domain—back when open reel tape served as the target preservation format—could be relatively simple, perhaps involving only an archivist and/or librarian and an audio engineer. Now that audio preservation—at least in terms of target format and long-term storage²—resides squarely in the digital domain, such a simple system is not possible. Audio preservation necessarily requires the contributions of not only the above-mentioned personnel, but digital librarians, IT specialists, computer programmers, intellectual property experts, content specialists, and others working collaboratively across disciplines within a relatively more complex system. To date, there have been few working models of successful systems of this type. We feel that the audio preservation field needs the following:

- Detailed and specific models of complete digital preservation systems
- Sets of best practices for successful operation of these systems. These must address in detail such topics as the definition and management of preservation digital files, selection

of items to be preserved, appropriate preservation signal chains, quality control, interim storage, long-term storage, data integrity over time, and many others

- Tools to enable functionality in many parts of the preservation system
- Tools that enable increased throughput and use of automation, while maintaining preservation-quality output, in order to preserve the overwhelming number of deteriorating audio recordings held in U.S. recorded sound collections
- Funding for the development of infrastructure and personnel to support and operate preservation systems

The *Sound Directions* project will provide solid examples of all of the above with a publication scheduled for release in spring, 2007 as well as through future phases of the project. However, due to the scarcity of solid models, best practices, and tools currently available to this field, we believe that additional work to further develop these is not only valuable, but necessary.

B. Access Systems

Recorded sound archives are charged with providing for both preservation and access of their holdings and, indeed, one does not hold much value without the other. There are currently few, if any, systems that can provide deep access to unique collections of field recordings along with their associated documentation. By contrast, there are a number of systems that provide good access to commercially-produced recordings.³ We feel that the audio preservation field needs the following:

- Proof of concept of an access system that can deliver field collections to researchers. This must include materials that make the recordings understandable such field notes, photographs, technical metadata, descriptive metadata, and navigational metadata, among other things
- A study of the discovery and delivery needs of researchers using field collections
- Actual development of a system as described above that is usable at more than one institution

C. Preservation Repositories

With the preservation of recorded sound collections now almost completely in the digital domain, issues relating to the management and long-term storage of digital data are of paramount importance to achieving long-term preservation. Many institutions are building preservation repositories, usually following the OAIS framework, to manage and store archival digital data in the form of preservation packages. However, the development of these repositories is in their infancy, with only a handful currently storing preservation packages for audio. Experience with preservation repositories in this country is limited, and this experience has largely not been tested by failure. Given that content providers intend to trust their data to such repositories, we feel that much research and development must be done in this area, exploring:

- The modeling of preservation repositories with various functionalities
- The creation of detailed and specific best practices for operating these repositories

- The modeling of service-level agreements between content providers and preservation repositories
- Long-term data integrity and authenticity issues
- Modeling of potential threats to content including such issues as obsolescence, failure of hardware and software, malicious behavior, human error, disaster and others, including the development of proposed solutions
- Migration strategies and tests

Some work in these areas is underway at various institutions, national organizations and in the digital library community, some with funding through the Library of Congress' National Digital Infrastructure and Preservation Program (NDIPP). Given the relatively long period of time it takes to bring preservation repositories on-line, the overwhelming amount of audio content in need of preservation, and the rapidity with which some audio formats are deteriorating, we believe that these efforts must be accelerated to provide for the long-term storage of endangered content before it is too late.

D. Smaller Institutions

The Archives of Traditional Music, while part of a large and well-funded academic institution, has limited resources and is pursuing the development of a preservation system that, with the aid of outside funding from granting agencies or private donors, is obtainable by smaller recorded sound collections. In the U.S. there is much audio content of regional and/or national importance that is held by small institutions with limited resources and technical expertise. While some of these institutions may be able to develop parts of a preservation system—from selection to the creation of preservation-worthy digital objects, for example—they are often unable to implement robust and enduring solutions for the long-term management and storage of their digital content. Although there is talk of small-scale mass storage systems using software that is both relatively inexpensive and requires less technical expertise, they are not yet on the horizon. If audio content held by smaller institutions is to survive, we believe that the following must be addressed:

- Modeling of mutually-beneficial partnerships between larger and smaller institutions
- Modeling of mutually-beneficial partnerships between smaller institutions and the private sector
- Development of regional or consortial models for the operation and support of preservation repositories
- Development of the concept of a transitional repository—an idea that originated with John Spencer of BMS/Chace—that safely holds content until a fully-functioning preservation repository is available.
- Development of best practices for a transitional repository including such topics as storage formats, data management and tracking, metadata collection, planned future migration, and many others
- Development of trustworthy commercial or academic repositories that will store content for a fee

The ARSC Technical Committee is beginning to explore some of these issues, but there is much more to be done.

E. Analog Playback Procedures

Both *Sound Directions* institutions have made use of experts in the reproduction of analog tapes and discs to train their audio engineers during the course of developing their preservation systems. We are keenly aware that this knowledge resides in fewer and fewer engineers with each passing year. Because digital recording technology and techniques now dominate the recording industry, audio engineers entering the field do not necessarily need, or possess, the knowledge and skills necessary to transfer analog recordings. Because recorded sound collections will need to play back their analog holdings for at least the next generation, if not longer, we believe that the following must be pursued:

- Document the experience, procedures, techniques, and knowledge of experts in analog playback. This documentation should not be high-level, but detailed and specific
- Make this documentation widely available to holders of recorded sound collections, audio engineering programs, and audio engineers who aspire to undertake preservation transfer work

F. Analog Playback Problems

There have been few scientific studies of the degradation mechanisms that afflict audio recordings. For some formats, such as lacquer discs, we must rely on work that was completed nearly 50 years ago.⁴ For other formats we know that our understanding is limited, and that we are often guessing about appropriate playback, restoration, and preservation procedures. For example, Richard Hess' work in consultation with a group of scientists, audio engineers, and tape specialists, has demonstrated that what has been termed loss of lubricant in open reel tapes is not truly this problem but deterioration caused by a number of factors not yet completely understood. His work also suggests that the mechanism by which baking (also called incubation) renders a Sticky Shed Syndrome tape playable has also been misunderstood. Hess has proposed a new category using the term Soft Binder Syndrome (SBS) for all polyester-backed tapes that exhibit sticking, squealing, and abnormal shedding.⁵ Important conceptual work such as this must be verified through actual testing to provide a solid understanding of how recordings deteriorate along with appropriate solutions for treating, restoring, and preserving them. If we continue 'shooting in the dark,' we will continue to lose content. We suggest pursuing an ambitious research agenda, including the following:

- Undertake a study on how polyester-based open reel tapes deteriorate, including recordings that fall into the Sticky Shed Syndrome category and those that don't
- Research the condition known as the Vinegar Syndrome that affects acetate-based open reel tapes. This problem is well-researched for film but has not yet been addressed for audio tape
- Revisit Pickett and Lemcoe's 1959 work on lacquer discs
- Conduct a study on degradation mechanisms that affect cylinder recordings. This study should develop recommendations for ideal environmental storage conditions for this format as well as procedures for cleaning these items

- Continue developing procedures and systems for non-contact playback of audio recordings, such as the collaborative work between the Library of Congress and scientist Carl Haber that is now underway

G. Training of Audio Engineers and Digital Librarians

Both *Sound Directions* institutions use professional audio engineers with knowledge of historical formats for preservation transfer work. The Archives of Traditional Music has extensive experience with both talented graduate students and audio engineers and understands that the critical listening skills and technical knowledge that resides in engineers is necessary for preservation-quality work. Technical expertise is also necessary on the repository side of a preservation system and this is typically found in digital librarians, information technology professionals, or a combination of both. We suggest pursuing the following:

- Encourage the creation of training programs for audio preservation engineers. Programs might be located in audio engineering schools or sponsored by any appropriate institution as an institute or workshop for engineers who already possess a degree or studio experience, and who want to learn this craft
- Encourage the development of audio-related curricula for library students specializing in preservation management, as well as the creation of training programs for audio preservation managers, who may come from an audio engineering, library science, or IT background and require further training in other areas
- Encourage programs that train professionals for work with digital libraries and preservation repositories, including metadata specialists, system analysts, hardware experts, and others

H. Promotion of Audio Preservation

Many institutions with recorded sound collections do not fully understand the fragility of these materials, their historical and/or cultural importance, or the measures that are needed to ensure their long-term preservation. Others simply do not prioritize these types of materials. We suggest the following:

- The creation of publications by NRPB and the Library of Congress that outline the problem and highlight the importance of audio preservation that can be used by recorded sound collections to leverage institutional support
- A national plan to promote audio preservation problems and needs that extends past the current study

Addressing any or all of the issues outlined above requires funding. The creation of an ongoing funding source for audio preservation might be the single most enduring action that can be taken to support efforts across the nation to preserve our audio heritage.

¹ FACET—the Field Audio Collection Evaluation Tool—will be available publicly in spring, 2007

² Preservation systems also operate in the analog domain in terms of analog playback, the quality of which critically impacts the output of the system.

³ The Variations system at Indiana University is one, as is Smithsonian Global Sound, various systems produced by private companies, and others.

⁴ The classic study of how lacquer discs deteriorate is from the Preservation and Storage of Sound Recordings, A.G. Pickett and M.M. Lemcoe. Library of Congress. Washington: 1959. Pages 15-24.

⁵Hess, Richard L. 2006b. Tape Degradation Factors and Predicting Tape Life. Paper presented at the 121st AES Convention, San Francisco, Calif., October 5-8, 2006.