This is a slightly modified version of the narrative submitted to NEH as part of the *Sound Directions* proposal. It has been updated in a few places to reflect both our increasing understanding of the worldwide context in which our project will take place and the recent publication of IASA TC-04, *Guidelines on the Production and Preservation of Digital Audio Objects.*

I. Significance of the Project

A. 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. The Indiana University Archives of Traditional Music (ATM) and the Archive of World Music (AWM) at Harvard University propose a joint technical archiving project—a collaborative research and development initiative with tangible end results—that will create best practices and test both emerging standards and existing practices for digital preservation.

The *Sound Directions* project 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, the Archive of World Music at Harvard University and the Archives of Traditional Music at Indiana University. Most of us are now approaching audio digitization in similar, deliberately cooperative ways. There are few published standards or best practices for audio preservation. Committees of the Audio Engineering Society and the International Association of Sound Archives (IASA) have written best practices for some parts of the audio digitization process. However, the analog to digital conversion process is not complete until safe and secure storage is attained and a way to insure readability over time is developed. In addition to developing best practices in a number of areas, the work proposed by the ATM and the AWM builds on collective experience and recent work on audio digitization in an important new way: it contributes the final step to the process--the creation of interoperable digital audio preservation packages, containing audio essence and metadata, following the OAIS model.. This is a step that has never been taken before for archival audio. Only when we can feel assured that we have new programs in place that ensure the survival of our threatened cultural heritage, can we reliably take advantage of the dramatic expansions of access that digitization and the Internet afford.

The development of best practices and standards in many areas, especially the production of interoperable audio preservation packages, is the essential and exciting next step to insure the preservation of our national heritage of fragile and deteriorating recordings. The ATM and the AWM are poised to lead the way forward into this new frontier of digital preservation and access by initiating a highly collaborative and consultative research and development process, the results of which will be widely disseminated.

With the *Sound Directions* project we will:

a) Develop best practices and test both emerging standards and existing practices for archival audio preservation and storage in the digital domain and report our findings back to the field;
b) Establish, at each university, programs for digital audio preservation that will enable us to continue this work into the future, and which will produce interoperable results. This is

groundbreaking work that is considered to be the next necessary step for true digital audio preservation and access;

    c) In the process, preserve critically endangered, highly valuable, unique field recordings of extraordinary national interest.

While best practices have been and are being developed for the initial digitization process, they do not exist in many areas of the preservation chain. We will develop best practices by testing different procedures and techniques in selected areas including: specifications for master preservation files from different analog sources, management of digital files including names, announcements and embedded material identifiers, down-sampling and creating derivatives, ingestion and storage of digital audio objects in digital library repositories, implementation of preservation services (including data integrity checking) for digital audio in digital library repositories, quality control and checking procedures, and the interchange and reading of preservation files constructed using METS in an Archival Information Package (an AIP, following the Open Archival Information System model) between institutions. In each of these areas, and others, we will report on procedures and techniques that produced the outcomes we were seeking to the quality desired, as well as procedures that did not work within both the ATM and AWM preservation systems and workflows. This will be the most comprehensive and detailed development of best practices to date, covering many critical areas in the preservation chain. It will also, as noted above, be the first time that interoperability has been achieved for digital audio preservation packages from two archival institutions. This will yield valuable data from two different operations and perspectives for use by other institutions designing audio preservation projects.

B. Interoperability
Simply put, if every institution's buckets of bits are different in character they are idiosyncratic--not interoperable--and true preservation has not occurred. Real preservation depends on the usability and readability of files over an extended period of time. In addition, should one institution fail, this type of interchange guarantees preservation by enabling any engineer to access preserved content.

Interoperable files depend on appropriate metadata to insure readability over time, and the development of best practices for the collection of metadata will be a critical part of this project. To digitize and store a recording so that it can be migrated and preserved, descriptive, administrative and technical metadata are essential in order to understand and interpret the digital object. Opaque digital objects are difficult if not impossible to preserve. The development of compatible Submission Information Packages (SIPs), as proposed in this project, lays the groundwork for defining what constitutes a preservation object. The standards for the SIP already developed at Harvard offer a good place to start. Developing these standards further at two different institutions is critical, and the process of submitting them to the scrutiny of other professional engineers and digital library experts will enable further refinements. The lead engineer at the AWM, David Ackerman, has not only fostered the development of Harvard's audio preservation efforts in this area, but is guiding the creation of technical metadata standards for audio internationally by leading the Audio Engineering Society's SC0306, the Working Group on Digital Library and Archive Systems.

Further, *Sound Directions* will demonstrate that it is possible for different institutions to work within their differing workflows and physical settings and still attain preservation through the production of interoperable results. Thus the information generated through this work will generalize to other institutions who want to use the project's innovations but cannot redesign their audio studios nor completely alter their staffing situations in order to do so. Working together, the ATM at Indiana and the AWM at Harvard will develop methods and best practices that are largely system-independent, that can be adopted by other institutions without overhauling their existing operations.

One important byproduct of our project will be the creation by grant programmers of tools for technical metadata capture, workflow management, ingestion of preservation files and the dissemination of

interchangeable preservation packages. These tools will be generalized and documented for release as open-source software.

C. Content and Access
 The recordings chosen as test cases for *Sound Directions* will be drawn from the rich, outstanding and unique ethnographic field collections of the Archives of Traditional Music at Indiana University and the Archive of World Music at Harvard University. A complete list of these materials appears in Appendix E. Field collections have been selected based on the following criteria: a) research and cultural value; b) preservation needs; and c) recording format (in order to test the transfer of a range of formats for this research and development project.) At AWM, selected collections include historic field recordings from Egypt, Iraq, Iran, Afghanistan, Pakistan and India, unique documents of cultural history from regions of tremendous interest to Americans today. At Indiana, selected collections include critically important cultural materials such as music of Iraqi Jews in Israel, music from pre-Taliban Afghanistan, music related to the world's longest-running civil war in Sudan, and African-American protest songs from the 1920s through the 1940s.

Sound Directions is conceived in two phases, the first of which is the subject of this proposal. While the focus of Phase 1 is research and development in areas critical to audio preservation, the project will also result in the preservation of the above collections along with the creation of basic access to these materials. Phase 2 of the project, which will require a follow-on grant, will emphasize access. In Phase 2 each institution will create on-line digital audio archives beginning with the collections selected for Phase 1. The ATM will build the Cultures in Conflict Digital Archive (CCDA), creating on-line access to the recorded heritage of peoples around the world whose cultural practices have been threatened or abolished as a result of conflict. The AWM at Harvard will create a digital archive using its rich historical collections of classical and folk music from Iran, Iraq, Pakistan and India. Both institutions will also pursue a program of "digital repatriation," making access copies available to nations and communities whose recordings we house. Critical preservation problems, however, must be solved before we can move to providing this extended access.

D. Why this Project Now?
Recent years have brought forth significant public concern about the value of unique audio collections and the pressing need to reformat them to insure their survival. In the United States, the Council on Library and Information Resources (CLIR) wrote, in its proposal to survey audio collections that "collections of recorded sound are an irreplaceable record of the history and creativity of the twentieth century."[1] The proposal notes that "awareness that our audio heritage is in peril has reached the highest levels of government, but the needs remain great."[2] Efforts to make progress addressing the problem include the summit-type meetings such as the federally funded conference Folk Heritage Collections in Crisis in December of 2000 http://www.clir.org/pubs/reports/pub96/contents.html, the Save Our Sounds project at the Library of Congress that ensued http://www.loc.gov/folklife/sos/, the CLIR survey of unique audio collections held in academic libraries http://www.clir.org/pubs/abstract/pub128abst.html and the Sound Savings symposium at the University of Texas in 2003 http://www.arl.org/preserv/sound_savings_proceedings/introduction.html. Numerous workshops flowing from these efforts were held at professional societies to address the inextricably linked issues of preservation and access to these recordings, making the point that there effectively **is** no access without preservation. Indiana and Harvard have been integrally involved in these conversations.

At this time, the International Association of Sound Archives, along with many sound archives around the world, have come to the conclusion that long-term preservation of information contained on analog media requires transfer to the digital domain.[3] Sweeping endeavors such as the Library of Congress' Audio-Visual Prototyping Project, and local ones including Harvard's Music from the Archive project, Indiana's

Cultures in Conflict Digital Archive, as well as the Mellon-funded Indiana University/University of Michigan digital archive of ethnographic video, are all proceeding along similar, mutually informed lines.

What is needed at this time is for these leaders in the field to move forward with the knowledge we have to develop more detailed and comprehensive best practices, test emerging standards and engage in the production of interoperable digital preservation packages.

E. *Sound Directions* and Other Projects in the United States
We have researched a number of audio digitization projects and published recommended practices, looking for efforts similar to ours. Although we found no working projects that were as comprehensive or detailed as what we are proposing, we did find three projects with which we share certain commonalities or to which we have looked for insight.

1. Audio-Visual Prototyping Project at the Library of Congress
http://www.loc.gov/rr/mopic/avprot/avprhome.html

The audio preservation staff at the AWM has been keenly aware of the development of the Culpepper facility at the Library of Congress. AWM has created preservation procedures that are in step with this planned facility. Carl Fleischhauer advised us initially[4], and both David Ackerman and Robin Wendler (from AWM and the Harvard University Library Office for Information Systems) were then invited to offer extensive comments on the metadata issues in the Culpepper plans. *Sound Directions* will make use of many ideas generated in the planning for Culpepper, instantiating them into an actual project while extending them into the domain of interoperable files, an area that LC has not yet addressed.

2. Digital Audio Archives Project at Johns Hopkins and Indiana Universities

Funded by IMLS, this collaboration between the Johns Hopkins University Libraries and the IU School of Music has undergone some changes and is still in the process of receiving institutional approval. It seeks to create a workflow management system prototype for digitizing audio, making use of simultaneous multiple transfers and audio segmentation software. Project materials consist of open reel tapes of music recitals that were recorded in a consistent manner and are predictable in terms of their content and sequence. Given the wild diversity of archival field recordings on formats such as lacquer discs and deteriorating open reel tapes, it is highly unlikely that ethnographic collections could make good use of this approach.

3. George Massenburg Metadata Project

Working with colleagues in the commercial recording industry, George Massenburg is endeavoring to persuade production teams for recording labels to incorporate uniform metadata in their digital products to insure readability of files into the future. Massenburg has indicated his interest in using the metadata standards produced by David Ackerman's Audio Engineering Society committee. Our project pairs nicely with this work as, between the two, similar preservation procedures will extend from popular commercial recordings through materials recorded in the field.

Note that there are a number of projects as well as on-going work in other parts of the world, particularly Europe and Australia, that have developed preservation systems and established solid practices for preserving audio in the digital domain. Some of this work is reflected in the new IASA Technical Committee document, IASA-TC04 Guidelines on the Production and Preservation of Digital Audio Objects, while other parts remain unpublished.

II. Background of Applicant

In CLIR's *Folk Heritage Collections in Crisi*s, sound preservation consultant Elizabeth Cohen writes, "the development of successful preservation strategies will require the cooperation of computer scientists, data storage experts, data distribution experts, fieldworkers, librarians, and folklorists."[5] Indiana and Harvard bring together a powerful combination of leaders in all of the above fields. Our specialists are in constant demand for consultations with other institutions. Each institution has preservation repositories built on mass digital data storage systems and extensively developed digital library programs led by recognized leaders in digital access and web delivery. Both Harvard and Indiana are charter members of Internet2, which provides the advanced networking that will deliver high quality digital audio to classrooms, conference rooms, and desktop computers around the world. Each institution features leaders in ethnographic fields including folklore, ethnomusicology and anthropology. We are well positioned to share the results of our work widely. In short, given the particular resources available to us at our institutions, and the extent to which we are actively engaged with other institutions and leaders in the sound archiving community, we are well positioned to take on this challenge.

A. Indiana University

The Archives of Traditional Music (ATM) http://www.indiana.edu/~libarchm/ 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. Proof of the ATM's leadership in this domain can be demonstrated through the numerous major grants it has received over the years, from federal agencies such as NEH and private foundations such as Mellon. The ATM has ample experience both in preservation work and increasingly in digital audio projects. In the 1980s, for example, the National Science Foundation funded the transfer of ATM's famous cylinder collection onto 1/4" open reel tape. In more recent years, ATM has made several forays into digitization. One project, funded by NEH, resulted in the interactive CD-ROM publication *Music and Culture of West Africa: The Straus Expedition* (Gibson and Reed 2002). With funding from the Institute of Museum and Library Services and the Library Services and Technologies Act (LSTA), and in collaboration with IU's Digital Library Program (DLP), ATM has created on-line access to Hoagy Carmichael materials http://www.dlib.indiana.edu/collections/hoagy/. At present, ATM and DLP are completing another collaborative project in which they are creating the Starr-Gennett Digital Archive with LSTA funding.

Recently, as ethnographic methods have begun increasingly to include video, the ATM has taken on the challenges of video preservation and access through the EVIA Digital Archive project (EVIADA), a $1.3 million project funded by the Mellon Foundation http://www.indiana.edu/~eviada/. This project, which also involves the DLP and participants from Harvard, has in many ways helped prepare IU for the *Sound Directions* project. In addition to further strengthening relationships and collaborative workflows between ATM and DLP, the EVIADA project has helped ATM establish strong working relationships with IU's mass storage system and has served as an initial test of the Fedora repository system.

Finally, most pertinent to the present proposal is ATM's Cultures in Conflict Digital Archive (CCDA) Pilot Project funded by IU's Center for the Study of Global Change. The pilot project's purpose was a) to develop and test procedures and infrastructure for digital audio preservation and access; b) to generate data on staff and resource needs, technical procedures, documentation procedures, work plan, transfer and processing time for various recording formats; and c) to clarify questions and needs to feed into the development of a full preservation system.

ATM's partner in this project, the IU Digital Library Program http://www.dlib.indiana.edu/, 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, the School of Library and Information Science, and the School of Informatics. The DLP's current facilities include the Digital Media and Image Center (containing equipment for image, audio, and video capture), the Library Electronic Text Resource Service, and an extensive server infrastructure for support of digital projects, with life-cycle replacement funding for hardware and software. DLP staff provide expertise in planning, creating and maintaining digital projects. DLP's Variations2 digital music library project http://variations2.indiana.edu/ 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. DLP Director Kristine Brancolini and Associate Director for Technology Jon Dunn have worked extensively with ATM staff, also participating in meetings with Harvard staff in the planning of *Sound Directions*.

The Massive Data Storage System http://storage.iu.edu/mdss.html is a distributed storage service offered by Indiana University's University Information Technology Services that is further described in the Methodology section below. This system consists of nearly 1.6 petabytes of disk and automated tape storage with files mirrored between servers at both the IU Bloomington and Indianapolis campuses.

B. Harvard University

The Archive of World Music and its technological partner, Harvard College Library Audio Preservation Services, are both units of the Loeb Music Library (http://hcl.harvard.edu/loebmusic/) which, in turn, is a component of the Harvard College Library that serves the Faculty of Arts and Sciences at Harvard. The Archive of World Music was established in 1976 and, with the appointment in 1992 of Kay Kaufman Shelemay as Harvard's first senior professor of ethnomusicology, the Archive moved to the Music Library to become one of its special collections. It is devoted to the acquisition of archival field recordings of musics world-wide as well as to commercial sound recordings, videos, and DVDs of ethnomusicological interest.

The Archive quickly attracted major collections including the James Rubin Collection of Indian Classical Music (probably the largest collection of Indian classical music in the U.S.), the Kay Kaufman Shelemay Collection of Ethiopic Musics, the Sema Vakf Collection of Turkish Classical Music (probably the largest outside of Turkey), and the Laura Boulton Collection of Byzantine and Eastern Orthodox Chant. Collection development has focused primarily on the Middle East, Asia (broadly understood) and Africa.

The Archive developed the Harvard College Library Audio Preservation Services (HCL APS), a state-of-the-art facility managed by an internationally known engineer. Over the past five years HCL APS has moved toward joining its counterpart, the HCL DIG (Harvard College Library Digital Imaging Group) in providing top quality service and advice for digitizing media. Both work closely with the Harvard University Library Office for Information Systems on matters of building robust infrastructure and sustainable tools for creating and preserving digital objects via the Digital Repository Service.

Substantial grant funding from The Laura Boulton Foundation, the Sema Vakf Foundation and the Harvard University Library Digital Initiative have provided funds for preservation and access to the AWM's collections and for building substantial infrastructure to support long-term digital preservation.

The AWM and HCL APS work in the context of an excellent overall preservation program centered in the Weissman Preservation Center (http://preserve.harvard.edu/hul/overview.html) directed by Jan Merrill-Oldham who recently won the 2004 Paul Banks & Carolyn Harris Preservation Award, given by the Association for Library Collections & Technical Services (ALCTS) in recognition of years of excellent

leadership in preservation.   Merrill-Oldham, with her high standards and vision, has served as a mentor and guide for the development of audio preservation at Harvard. Together we form a leading national preservation program of recognized accomplishment.

The Harvard University Office for Information Systems (http://hul.harvard.edu/ois/) coordinates all of the Library's online catalogs (HOLLIS, its MARC catalog, OASIS for finding aids, VIA for visual images, and so forth) as well as the highly regarded Library Digital Initiative (LDI), the Digital Repository Service, and innumerable tools that sustain and support online resources. Led by Dale Flecker and Tracey Robinson, OIS is home to nationally recognized experts such as Stephen Abrams and Robin Wendler, who will advise the current project. The Library Digital Initiative in some aspects parallels IU's Digital Library Program. Its mandate is to create the technical infrastructure to support the acquisition, organization, delivery, and archiving of digital library materials, provide experts to advise the community on key issues in the digital environment and enrich the Harvard University Library collections with a significant set of digital resources.

 The AWM and HCL APS have years of successful experience working with the Library Digital Initiative and OIS. Together we have created significant infrastructure in support of audio preservation including Dmart, a tool for uploading audio files and attendant metadata into the Digital Repository, and an audio processing XML editor (APXE) for the efficient collection of audio metadata.

III. Project History

A. Overview
The *Sound Directions* project was born in March 2003 when Daniel Reed and Virginia Danielson were serving on the Council on Library and Information Resources committee to create a survey tool for academic libraries with audio collections. During a break, Reed and Danielson began discussing plans at their respective archives, realized they were heading in similar directions and immediately saw the potential advantages of collaboration. Points in common included:

- both were conducting digital-only pilot projects;
- both were convinced of the need to move toward digital preservation;
- both wanted to increase web access to collections;
- both had all the resources to accomplish the above—highly valuable, unique collections with critical preservation needs; mass digital storage; exceptional personnel including national leaders in critically important areas (e.g., digital library staff, archival audio engineers).

Over a year of planning has ensued, which included three face-to-face planning meetings between Reed and Danielson: at the Society for Ethnomusicology conference in Miami in October 2003, at Harvard in November 2003 (also involving Mike Casey and David Ackerman), and finally at Indiana University in March 2004 (involving librarian Sarah Adams of Harvard, David Ackerman, the entire staff of the ATM, and representatives of IU's Digital Library Program, Variations 2 project, Massive Data Storage Service, and the EVIA Digital Archive). In these meetings, and through consistent phone and email dialog, this project began to crystallize. A critical point in the process was when we realized the extent to which emerging standards remain untested in real projects, how little information is available in terms of best practices for digital audio preservation and the importance of producing interoperable files.

B. Pilot Projects
The pilot projects that the ATM and the AWM have pursued have provided us with critical background experience and information that inform this proposal. ATM's Cultures in Conflict Digital Archive (CCDA) Pilot Project was designed as a limited test run through digital preservation and access in the IU environment. Through this project, ATM made use of the resources presently at hand in order to clarify

questions and needs that could be addressed in a fully realized, research and development project of the kind we are currently proposing. We evaluated every aspect of the process, from staff and equipment needs to IU's digital infrastructure to work flow. The CCDA pilot project helped us identify what emerging standards beg to be tested, and clarify in what areas we might pursue a more fully developed study of recommended best practices.

The Archive of World Music's pilot project, "Music from the Archive: A New Model of Access to Rare and Unique Sound Recordings," received a grant of $204,000 from the University's Library Digital Initiative Program. Focused on three of the AWM's premier collections, "Music from the Archive" integrated musical sound and related images in the same electronic research tool (see, for instance, http://oasis.harvard.edu/html/mus00001frames.html). The project developed technologies for access and digital preservation of rare materials and it advanced ways of integrating digital access resources and digital objects, providing a knowledge base that will inform the present project. "Music from the Archive" produced the tools that form the infrastructure of the AWM's (and possibly the ATM's) work on the present project: we adapted EAD finding aid protocols to create, for the first time, multi-media finding aids for ethnographic collections; we contributed substantially to the development of standards for audio metadata, and created software for the collection of metadata and for the deposit of audio files in the Digital Repository. We also created an internal Submission Information Package for the deposit of archival audio in the Repository. This will serve as a starting point for work on the project at hand.

C. Duration
Phase 1 of the *Sound Directions* project is by itself a self-contained, cohesive project that will be 18 months in duration.

IV. Standards and Best Practices

A. Introduction
The *Sound Directions* project will research and develop best practices and test emerging standards in many critical areas of the preservation chain including the gathering of technical metadata, construction and exchange of preservation packages, creation of derivatives, quality control, workflow, ingestion and storage in a preservation repository, and the construction of digital files for archival preservation. We will develop many best practices in these areas including:

- Use of tools for building SIPs using the METS schema;
- Creation of interoperable preservation packages;
- Use of tools for collecting technical audio metadata;
- Use of appropriate file formats for archival audio preservation including what should be part of the file;
- Amount and type of redundancy desirable for archival audio preservation;
- Who should perform quality control and what procedures should be used.

There are many additional best practices that we will research and develop. These are discussed in detail in the paragraphs below and summarized in the table in Appendix A.

B. Overview of Standards and Best Practices
Formal standards in preservation-related areas are assessed and ratified by bodies such as the International Organization for Standardization (ISO), the National Information Standards Organization (NISO), the World Wide Web Consortium (W3C) and others. Standards crucial to audio preservation are also developed by organizations such as the Audio Engineering Society (AES), the European Broadcasting Union (EBU), the Library of Congress, the Digital Library Federation (DLF) and others that may not be official national or international standards organizations in the strictest sense, but are charged by various

constituencies with providing leadership in this area. Indiana and Harvard, like other academic institutions, are not formal standards-setting bodies. We are well-positioned, however, to make significant contributions to emerging standards. We are also positioned to test, research and develop best practices for archival audio preservation.

With this technical archiving project, the Archives of Traditional Music and the Archive of World Music will:

- Research and develop best practices in many key areas where standards currently do not exist;
- Research and develop best practices necessary to meet existing and emerging standards;
- Test existing and emerging standards with a real world project;
- Provide extensive data derived from this actual project on what works and what does not in many parts of the preservation chain.

The results of our work on this project will:

- Provide institutions wishing to develop audio digitization preservation projects with tested best practices for many links in the preservation chain;
- Help shape the development of future standards;
- Present strategies and procedures for meeting the requirements of existing/emerging standards, some of which have not have been tested through multiple, real world projects;
- Document problem areas within existing/emerging standards.

C. Standards
Few formally-ratified standards exist at the current time for the preservation of audio recordings in the digital domain by archival institutions. The Open Archival Information System (OAIS) was accepted by the ISO in 2002-03 as a reference model addressing a full range of archival information preservation functions including ingest, archival storage, data management, access, and dissemination (ISO 14721:2003). Although not specifically an audio standard, the model is becoming widely accepted by institutions undertaking a broad range of preservation work. Emerging standards include the Metadata Encoding and Transmission Standard (METS) maintained by the Library of Congress while being developed by the Digital Library Federation. They also include on-going work by AES to develop methods and schema for tracking technical audio metadata, as well as work done by the EBU to develop the Broadcast Wave File format.

The *Sound Directions* project will serve as a test site for the following existing and emerging standards:

1. AES-31-3-1999 http://www.aes.org/publications/standards/

This file exchange standard, published by AES in 1999, is used to model the relationship between the source recording and resulting digital files. Although not developed with archival preservation work specifically in mind, the standard can apply to transfer situations with deteriorating recordings that require multiple stops and starts by the engineer to obtain full playback. It provides a way to link the various files that are created, thereby reconstructing the source recording. Without it, future researchers are left with one engineer's interpretation of the edit points. In this project we will work on defining situations that we feel require the use of the standard while examining places where the choice might be less clear. We will then develop procedures for implementing the standard within each of our workflows. We are not aware of any audio preservation projects undertaken by archives other than the Archive of World Music at Harvard University that have employed this standard.

2. Audio Engineering Society SC-03-06 Working Group on Digital Library and Archive Systems, Task Group SC-03-06-A Metadata Harmonization

These emerging standards, developed in consultation with the Library of Congress by the AES in a working group chaired by Harvard's David Ackerman, are used to gather technical metadata on both the source recording and the digitizing process. They set forth the metadata elements and schema used to describe the basic characteristics of audio digital objects from their creation through numerous migrations. One standard enables detailed documentation of the source recording and its digitized derivatives. The second captures data on the digitization process, essentially allowing the taking of a snapshot of a studio setup through each step of audio capture and transformation. Both of these emerging standards are nearing the end of their development cycles but have yet to be implemented in a real world project. We will develop procedures, strategies and best practices to enable implementation of these two standards. This will include the development of computer-based tools and databases for collecting and tracking technical metadata.

3. Broadcast Wave File Format (BWF) http://www.ebu.ch/trev_274-chalmers.pdf

The Broadcast Wave File format, based on the Microsoft WAVE audio file format, was introduced by the EBU in 1996 to allow files to be exchanged between the increasing number of digital audio workstations used in radio and television production. The format was developed in collaboration with the industry and has been implemented in equipment from a large number of manufacturers. BWF is a special type of WAVE file that may contain basic, though limited, metadata about the audio content that resides with the file itself. Although this is not a destination for the extensive metadata that must be collected during digital preservation projects, the format itself is rapidly becoming a standard in the audio world. In addition to its widespread use in Europe and other parts of the world (it is recommended by IASA), the AES and the National Academy of Recording Arts and Sciences have specified its use for archiving purposes in this country. BWF has been implemented by a few archival institutions undertaking digital audio preservation projects in the US, mostly in a digital audio workstation environment. In this project, we will explore implementing BWF in both a native and dsp-assisted digital audio workstation environment while examining its role in audio preservation projects that are already using an extensive metadata system.

4. Open Archival Information System (OAIS)
http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf

The Open Archival Information System (OAIS) Reference Model is a conceptual framework for an archival system dedicated to preserving and maintaining access to digital information over the long term. It describes the environment in which an archive resides, the functional components of the archive itself, and the information infrastructure supporting the archive's processes. OAIS was originally developed by the Consultative Committee for Space Data Systems in support of space research data, but in the process of ISO standardization, the model was reviewed by a larger community including libraries and archives. Due in part to endorsement by OCLC and RLG (Research Libraries Group) (reference: http://www.oclc.org/research/projects/pmwg/pm_framework.pdf), the OAIS Reference Model is being used by many libraries, archives, and other cultural heritage institutions as a means of defining their own digital preservation infrastructure. Indiana and Harvard are using OAIS concepts in implementing their digital library object repository systems.

5. Metadata Encoding and Transmission Standard (METS) http://www.loc.gov/standards/mets/

The Metadata Encoding and Transmission Standard (METS) specifies an XML document format for encoding metadata necessary for both management of digital library objects within a repository and

exchange of such objects between repositories, or between repositories and their users. A METS document is capable of packaging together all of the descriptive, administrative, and structural metadata for a digital object, plus references to the object's data files, or optionally, the data files themselves. METS has been designed with the OAIS Reference Model in mind, such that a METS document could play the role of Submission Information Package (SIP), Archival Information Package (AIP), or Dissemination Information Package (DIP). METS is an extremely flexible format, supporting the inclusion of metadata in nearly any format for which an XML schema exists. A system of "profiles" allows for the definition of particular combinations of metadata schema for particular applications, such as digital audio archiving. We will explore the suitability of METS as a means of relating the many digital objects that make up the digital audio representation of an original audio object and the metadata related to those objects, in the context of the SIP/DIP. We seek to establish a METS implementation that can be used to facilitate exchange of archived digital audio between institutions.

D. Best Practices

The publication of best or recommended practices provides guidance in areas where standards do not yet exist or may never be created. Best practices may also provide strategies, procedures or work plans necessary to successfully implement a standard that has been formally adopted. There are few published sets of best practices that address the preservation of audio recordings in the digital domain.[6] Of those that exist, none that have emerged from actual projects are comprehensive or detailed enough to meaningfully assist archival institutions holding audio materials that are building preservation systems.[7] We have discovered the following problems with the published best practices we consulted:

- Some, while helpful, are not based on actual projects but reflect the current thinking of individuals or institutions (see endnote 6 for a discussion of the forthcoming IASA TC04 document);
- Many are too general to be useful. They do not provide enough detail to inform decision-making in areas that are crucial to the success of a project;
- Some are focused on digitizing for access, not preservation; some address building digital libraries, not digital archives;
- None are comprehensive enough to be of deep value, omitting topics that must be addressed for a successful and enduring project.

It may not be surprising that, at this early stage in an emerging field, published best practices are not comprehensive. Eventually we might hope to find practices that cover all parts of the audio preservation chain from cleaning and analog playback to file storage and interchange between institutions. Each of these areas, and points in-between, must be addressed in detail for enduring preservation to be achieved. The development of comprehensive best practices in an area with this amount of complexity will necessarily unfold over many years using data from a number of institutions and individuals. This effort will benefit from the completion of multiple projects, even if they address some similar issues. Indeed, the Library of Congress, working with the Council on Library and Information Resources (CLIR), has begun a long-term study involving a number of individuals and organizations that will strengthen the development of best practices in this area.

One of the primary goals of the *Sound Directions* project is to research and develop best practices in a number of critical areas where they do not yet exist, to a level of specificity not yet seen, contributing this data to the on-going national effort. We will not duplicate work that is already underway, although we may have findings or opinions from our project in these areas that we will put forward. Below is a list of critical links in the audio preservation chain along with a discussion of what we expect to achieve in each area.

1. Physical restoration of analog source recordings: Physical restoration refers to the process of preparing a recording for optimal playback. It may include cleaning or repairing discs, baking tapes, dealing with fungal growth, etc. This is a much under-studied and controversial area in which there are a variety of opinions and techniques. The engineers at ATM and AWM have a number of procedures that they have learned and developed over many years. Although it is beyond the scope of this project to research physical restoration techniques, they will report on the efficacy of these techniques within this project.

2. Analog playback of source recordings: This area includes such topics as appropriate classes of equipment for playback, condition and maintenance of equipment and playback techniques and strategies. The LC/CLIR study is focusing on this topic so we will not duplicate this work. Also, the new IASA document cited above explores this area in some detail. Again, the engineers at ATM and AWM have years of experience with achieving optimal analog playback and will report on their experiences in this project.

3. Analog to digital conversion and sampling rate/bit depth choices: This is another area being explored by the LC/CLIR study and addressed in the IASA document. Despite the recommendation of some published best practices there is a general consensus that, for preservation work, a good outboard analog to digital converter should be used to create high-resolution files. Both ATM and AWM have excellent converters and much experience creating high resolution (24/88.2 or 24/96) files. We will report on what we use and how it works,..

4. File formats and characteristics: This area is fertile ground for the development of best practices. None of the published documents we consulted provide significant detail in this area. Most of them suggest preserving a flat, unaltered or raw version of the audio content but none define specifically what this means. We will undertake research and development in this area, addressing the following questions:

- What role might the Broadcast Wave File format play in a preservation project where extensive metadata is already collected and stored? Are there enough significant advantages to using BWF in this context that its use should be specified?
- How should files be named? What role might names based on persistent identifiers play? How much, if any, content should be carried in a file name?
- Are voice announcements that are placed in the file, residing directly with the audio essence, useful? If yes, should these announcements carry technical information? Content or numbering information? Signify the end of content?
- Should markers and regions that mark divisions of the content of the audio recording be included in the file? What resources (time) and expertise are needed to divide content in this manner? How should the location of these markers be included in the metadata system?
- What characteristics define a flat or unaltered version of the audio content?
- When are signal-processed versions of the audio content appropriate? Not appropriate? What resources—time and expertise—are needed to signal process? Is there an approach or philosophy to signal processing that is useful in generating versions for archival research?

5. Technical Metadata: As noted under standards above, we will develop best practices and produce tools and databases necessary for implementing two emerging technical metadata standards generated by AES. This is uncharted territory: although some institutions are gathering metadata elements that may overlap with the AES set, there has yet to be a project that actually implements these standards. In pursuing this research and development we will address the following questions:

- Who is best able to gather technical metadata?
- What resources—time and expertise—are necessary?

- At what stage in the preservation process (or a project's workflow) should the various types of technical metadata be collected?
- What procedures enable the most efficient uses of the metadata collection tools?

6. Copies (Derivatives): Decisions on the type of derivatives to produce from a preservation project are driven largely by an institution's access needs and technical capabilities. Although this is institution-specific, there are issues relating to the process of producing derivatives that are ripe for exploration. There are also preservation issues around questions of redundancy that have not been addressed in detail in previous best practices. We will undertake research and development in this area, addressing the following questions:

- Who should produce derivatives? What resources—time and expertise—are needed?
- What technical issues impact quality in the way in which derivatives are produced?
- What amount/type of redundancy is useful for an archival preservation project in the digital domain? What kinds of derivatives can fulfill redundancy requirements?
- How durable should derivatives be? Is there a relationship between access goals and durability?

7. Storage: Many institutions have implemented or are starting to implement digital preservation repository systems to support the long-term storage and management of digital content. In some cases, such systems are being based on commercial digital asset management software, while in others, open source repository software such as Fedora or Dspace provides the foundation. Due to the large file sizes involved, digital repositories often integrate with hierarchical storage management (HSM) systems based on automated tape libraries.

Repositories require standards for ingestion, storage, and dissemination of digital objects and metadata, and both AWM and ATM have implemented or are implementing digital repositories which make use of METS for this purpose. Repositories also must provide certain services to support long-term storage of digital content, including file integrity checking and migration of content due to media or file format obsolescence. We will tackle several research questions in this area, including:

- What are the necessary components of a SIP, AIP, and DIP in the OAIS model for digital audio preservation?
- How can these packages be represented, most likely taking the form of a METS profile?
- How can these packages be stored in at least two different digital repository systems (Fedora at IU and the DRS at Harvard)?
- Can packages be exchanged between the two repository systems?
- How are preservation repository services such as data integrity checking best implemented in an HSM environment?
- Based on current and expected storage technology, what migration cycle will be necessary in order to keep up with media changes?
- How can archivists best work with IT staff to develop and implement a migration strategy?
- What is appropriate interim, on-site storage for files waiting to be uploaded to the preservation repository?

8. Quality control: If quality control issues are addressed at all in published best practices they are usually not explored in any detail, yet they are a critical component to any archival preservation project.[8] Checking systems may be implemented to assess the work produced by project engineers and other staff, the physical media to which derivatives are copied, and the copied derivatives themselves. In conducting research and development in this area we will address the following questions:

- Who should perform quality control? Can quality control in different areas be performed by different people? Who checks digital files? Collected metadata? Derivatives?
- What resources—time and expertise—are needed?
- What specific procedures should be followed in a quality control process?
- What level of checking is enough? What level is too much?
- Can the checking of digital files be automated, using known file characteristics and gathered metadata as a baseline?

9. Exchange of preservation packages: One of the best tests that enduring preservation has been achieved is the interchange and reading of preservation files between institutions. This type of interchange guarantees preservation, assuring the viability of preserved content should one institution fail by enabling any engineer to easily access another's work. To our knowledge, this has never been done by archival institutions engaged in audio preservation projects in the digital domain. In this project we will research and develop procedures and strategies for interchange, resulting in a set of best practices. Research and development questions in this area will include:

- Can we successfully demonstrate interchange of metadata and content between two institutions?
- By what criteria should success of interchange be determined?
- What shared technical infrastructure is needed for interchange? What elements can be different?
- What components of a SIP/DIP are necessary for interchange? What components are optional?
- Are there procedures or practices that make interchange more feasible?
- What changes are necessary to the IU and Harvard digital repositories to enable interchange?

10. Workflow: This area is rarely addressed in audio best practice documents yet it is also essential to a successful and efficient preservation project. Most archives hold significantly more recordings than they can realistically preserve. Designing workflows that address quality issues, while striving towards efficiency and productivity, enable a larger number of recordings to be preserved. In this project, ATM and AWM will test two different workflows developed for archival preservation of deteriorating sources as they reach for common goals. In conducting research and development in this area we will address the following questions:

- Who does what? What expertise is needed in each link of the audio preservation chain?
- How much time is required, on average, to complete each step in a preservation project?
- At what points do issues relating to quality intersect with personnel, tasks and time?
- Are there management procedures/systems that are useful in tracking workflow, organizing tasks and encouraging efficiency?
- What technical infrastructure is needed to support a productive workflow?

V. Methodology

A. Collaboration
This project pairs one of the country's major public academic institutions with one of its premier private universities, working in different settings towards shared outcomes and goals. Each institution brings to the project different perspectives on how to reach these goals, resulting in preservation systems that address the same problems from different angles. Both institutions have staff with many years of experience reproducing analog recordings to the highest quality possible using professional analog playback devices. These staff members have successfully completed a number of analog preservation projects over the past ten years and have also completed pilot digital projects. Our differing approaches to preservation in the digital domain will enable this project to provide two models for achieving true preservation, widening its applicability to any number of other institutions and projects.

To successfully accomplish the research and development objectives outlined in the section on standards and best practices above, ATM and AWM will work collaboratively with both existing and newly-hired project staff from each institution. Our shared overall goal is to produce preservation packages that can be ingested into, and then output by, each other's preservation repository. To insure that the work we do to reach this goal is both high quality and enduring, we have identified the following general principles to guide us:

- Use existing or emerging standards where available;
- Consult with each other and with outside experts in digital audio preservation and digital library development;
- Use the following question as the primary test for potential solutions: does the proposed solution lead to high-quality, replicable and enduring results?

To further guide our work we have also adopted the recommended practices and strategies outlined by the technical committee of the International Association of Sound Archives in their document *Standards, Recommended Practices and Strategies* (IASA-TC 03).[9] While this document does not present formally-adopted standards, nor is it based on an actual project, it does provide an excellent (though general) set of strategies for both playback of analog sources and preservation in the digital domain. Examining this document led us to formulate the following fundamental guidelines for our project:

- Employ professional audio engineers for all digitization work. All transfers will be attended;
- Conduct all transfer work in an audio studio that functions as a critical listening space;
- Use professional-quality playback machines and detailed, format-specific playback techniques for each type of recording;
- Use professional-quality outboard analog to digital converters;
- Create high-resolution files that preserve unaltered versions of the source recordings;
- Collect extensive metadata about the source recording, its derivatives, and the digitizing process;
- Store files in a managed preservation repository.

In fact, both ATM and AWM have already instantiated most of the IASA recommended strategies into their working procedures[10].

In addition to our overall goal of interchange, we share objectives in each of the areas in which we will research and develop best practices. With quality control, for example, one of our objectives is to develop workable and consistent procedures for the checking process. While our objective is shared, the specific end result (the actual procedures), and the process of arriving at the result, may differ. Since we are implementing different workflows we expect our processes and, in some cases, our end results to diverge. In all cases our basic objectives and guiding principles remain shared.

IU and Harvard will work closely with each other throughout the project, identifying issues, devising strategies and sharing results obtained within each of our workflows. Project staff from the ATM and the AWM will hold face-to-face working meetings twice during the grant period, once in Washington DC and once in Bloomington. The first meeting will be dedicated to reviewing and discussing all technical issues related to project objectives. At this meeting we will also consult with each other to determine what software will need to be adapted or newly produced to support the work. The second meeting will allow us to measure our progress, address outstanding problems and change course, as necessary.

To further our collaborative work, and to insure that our results are broadly applicable to the field, we have formed an Advisory Board of technical experts that we will convene twice during the grant period to provide feedback and outside perspectives. Members of the Board are listed in the staff section below. At the beginning of the grant, ATM and AWM will present detailed proposed work plans/preservation

systems to the Board for feedback at a day-long meeting that will take place at the CLIR offices in Washington DC. This meeting will include an in-depth discussion of each part of the project. We will then build a feedback loop into the project by reconvening the Board midway through the grant period. This will enable further discussion of problematic areas as well as results achieved up to that point.

B. Indiana University
To meet project goals and objectives, IU will hire an audio engineer, a project assistant, and a programmer/database specialist. IU will also purchase equipment to meet technical objectives and guiding principles. Our working methodology is to explore preservation transfer in a native digital audio workstation environment using a professional audio recording/editing program with excellent outboard analog to digital and digital to analog converters. We will use a digital object repository management system, based on Fedora, along with the IU's mass storage system, to create a preservation repository.

The project will unfold at IU in two departments: the Archives of Traditional Music and the Digital Library Program (DLP). The ATM will be the primary location, providing project recordings and hosting all activities relating to digitization. DLP will handle programming and database work for the collection of metadata and for interfacing with our preservation repository.

1. Archives of Traditional Music

The project audio engineer will be supervised at the Archives of Traditional Music by the Coordinator of Recording Services (CRS), Mike Casey. Together, the engineer and the CRS will be responsible for Indiana's research and development in all areas relating to preservation transfer, including quality control and workflow. The CRS will devote 30% of his time to the project and will serve as the day-to-day manager for the Indiana portion of the grant.

The ATM will explore a workflow that pairs a project assistant with the engineer, dividing tasks from various parts of the preservation chain among them. The engineer will be responsible for optimizing analog playback, performing transfers, gathering technical metadata, performing signal processing if used, and creating or overseeing the creation of derivatives. The project assistant will gather, prepare and deliver collections and documentation to the engineer and manage workflow into and out of the studio, mark divisions of content in files and collect metadata, construct ATM indexes, check re-sampled files and technical metadata, produce physical copies and interface with the ATM Archivist and Librarian. Workflow at the ATM will also involve the ATM Archivist, who will have overall responsibility for the documentation (indexes) associated with project collections. The ATM Librarian will be responsible for creating or updating MARC format catalog records for project collections. Finally, the ATM Director will serve as principal investigator with overall responsibility for the Indiana side of the project.

2. Digital Library Program

The programmer/database specialist will be supervised in the Digital Library Program by Jon Dunn, Associate Director for Technology. This position will be responsible for the design and implementation of tools for: workflow management, technical metadata capture, and the creation, ingestion, and dissemination of interchangeable preservation packages. Also, investigating the use of a Fedora-based digital repository system for digital audio preservation, and generalizing and documenting tools developed on this project for potential release as open-source software.

3. Technical Infrastructure and Knowledge at Indiana

Technical infrastructure to support the proposed workflow and to tackle project objectives is largely in place, although some upgrades and additions are needed as explained below. Technical knowledge is also in place, as evidenced in the section on staff below. Indiana has the following capabilities:

a. Preservation Studio: Analog playback machines including Studer A810's, Otari MX5050's, Technics SP15's, Tascam 122MKIII's. Mytek Stereo96 ADC and DAC, and Apogee PSX100 converters; Mixing board and monitors both slated for upgrade as noted below; Dell Optiplex260 PC's, RME Digi96 PAD sound card, Plextor CD burner; WaveLab, Sound Forge, CD Architect, Waves Mastering and Restoration bundles;

b. Preservation Repository: Through major funding from Indiana University's Information Technology Strategic Plan, the Digital Library Program is currently beginning a project to implement a digital repository system to support access to and long-term storage of digital library objects in a wide variety of formats. This system will be based on the Fedora digital object repository system developed by the Cornell University and the University of Virginia with support from the Mellon Foundation, and will be a centrally-managed and centrally-funded resource for use by libraries, archives, and other units with digital collections across the IU system. This repository system will make use of IU's Massive Data Storage Service, a distributed storage service based on a consortially-developed software product known as High Performance Storage System operated by IU's University Information Technology Services. This disk- and tape-based system currently provides a total uncompressed data storage capacity of nearly 1.6 petabytes, and through use of the State of Indiana's I-Light optical fiber network, can support automatic mirroring of data between servers on IU's Bloomington and Indianapolis campuses for added disaster tolerance.

4. Technical Upgrades and Additions

To meet technical objectives and guiding principles we plan to upgrade our largest studio, designating it as a digital preservation studio. This space will be used full-time for the grant and, post-grant, for continuing digital preservation work. For this reason we will need to reconfigure existing space to create an additional studio space to meet our on-going work. We will also create a simpler quality control, file management and CD/DVD burning station that will be used to complete grant work, enabling the preservation studio to be dedicated to digitizing. IU will contribute equipment including such items as a passive switcher, patchbay, monitors, sound cards, and software. We will also upgrade several of our playback machines for grant work by installing additional playback heads. In addition, IU will purchase computers for the preservation studio and the checking station and install RAID 1 systems in both.

C. Harvard University
The locus for this project will be the Archive of World Music at Harvard, a division of the Loeb Music Library in the Harvard College Library.

1. Staffing and Responsibilities

The project will be managed by a team of three professionals experienced with the requirements of complex technological projects: David Ackerman (lead audio preservation engineer), Virginia Danielson (Richard F. French Librarian and Curator of the Archive of World Music) and Sarah Adams (Keeper of the Isham Memorial Library). Ackerman will guide the project, contributing 30% of his time to framing the questions, planning and supervising the work, working with the Advisory Board and Mike Casey at the ATM, and creating the reports for the project. Danielson will be responsible for overseeing the timetable of the project, administering AWM's share of the budget, and adjudicating any content issues that may arise. Adams will be responsible for intellectual

access to the project's products, insuring that EAD finding aids and MARC catalog records for the project's content are up to international standards. These three administered Harvard's Music from the Archive project and work well and efficiently together. The management team will be aided by Constance Mayer, experienced in issues related to user interface for electronic products, and Candice Feldt, our senior music cataloger.

From the Harvard University Office of Information Systems, Stephen Abrams (Digital Library Program Manager) and Robin Wendler (Metadata Analyst) will contribute advice and assistance in the software development that will be necessary and in the processes needed to enable the interoperability of files between the two institutions.

The AWM will hire three positions: 1 FTE audio engineer, .5 FTE programmer, and .3 FTE project assistant. The project engineer, under Ackerman's supervision, will carry out digital transfer, collect metadata, prepare notes for Ackerman's reports, and create derivatives. The project programmer, also under Ackerman's supervision and in cooperation with Abrams, will prepare software necessary to the work of the project. The project assistant will be Donna Guerra, the curatorial assistant in the Archive of World Music, who will add hours to her current part-time schedule with responsibilities for overseeing the conveyance of collections to and from the audio studio, creating those intellectual access tools that may not already exist and adding links from the finding aids or catalog records to digital objects produced by the project.

2. Overview of Project Timing

Initially, Ackerman will be keenly involved in planning for the first meeting of the Advisory Board. Because AWM has already invested in the development of a SIP as well as other tools for digital audio transfer and storage, it will be desirable for us to hire the project audio engineer as soon after that meeting as possible. Ackerman will advertise and hire for the position, train the successful candidate and frame the hands-on work that this individual will do. The relationship of Ackerman leading the project and the engineer as hands-on operator and information-gatherer should persist for the length of the project. At the first Advisory Board meeting discussions will be held by ATM and AWM participants designed to assess their state of software needs. Following that discussion, Ackerman will advertise and hire a programmer.

The project management team will already be in place and ready to go as will the project assistant. At present, an EAD finding aid is complete for one of our selected collections and we have begun work on a second, as well as catalog records for two more. At the start of the project, one additional finding aid will be needed and we anticipate that some tweaking of our completed work will be necessary. Donna Guerra, an experienced and independent worker, will be able to begin this work immediately. Adams will review Guerra's work and act as the project's liaison with the HOLLIS and OASIS online catalogs of MARC records and EAD finding aids.

3. Technical Infrastructure

The Audio Studio: As presently constituted, Harvard College Library Audio Preservation Services, featuring two audio preservation studios, is equipped to handle quarter inch analog tape in a variety of track formats, a wide range of audio disc formats, cassette recording and digital formats such as DAT, CD, DVD-Video and DVD-Audio. The studio includes professional converters that enable accurate digitization and monitoring of the audio materials without additional coloration. Studio A features a tuned control room with Genelec S-30D speakers, providing an excellent monitoring environment. For this project, the Harvard College Library will contribute a second Pyramix workstation equipped with a post pack, DDP Option, Sonic Solution Import, cue sheet support, Retouch plugin, Mykerinos DSP board, AES/EBU daughter card, and CD burner, Prism d/a and a/d converters and a pair of Genelec monitors. For additional details of studio equipment see Appendix H.

The Digital Repository Service: The Harvard University Library has established a managed digital storage system that will be used for this project. The Digital Repository Service (DRS) offers a storage and retrieval system that provides for the management of administrative and structural metadata

associated with stored objects. The DRS operates based on established preservation policies and procedures to ensure the continued usability of stored objects and delivery of these objects to software applications such as library catalogs. Additionally, the Office for Information Systems, working with David Ackerman, has nearly completed development of a METS archive tool for audio deposit (http://hul.harvard.edu/ois/systems/drs/dmart/current/). Working with a contract programmer, Ackerman has developed an audio processing XML editor (APXE) for the efficient collection of audio metadata.

D. Evaluation
The evaluation process for the *Sound Directions* project will make use of both formative and summative methods, combining the strength of an in-process with an end-of-project evaluation.

Formative methods of evaluation provide a way of periodically monitoring the progress of a project, assessing the quality and quantity of on-going work to insure that objectives are met in a timely manner. We will divide the project into three month development cycles and produce a brief evaluation at the end of each cycle. This will provide us with a tool to monitor the progress of the project at regular intervals, and provide for the tracking and addressing of problem areas encountered by the approximately twelve project staff located in four departments at two institutions. Five interim evaluations will be produced— the final report will cover the sixth cycle. These interim evaluations will be produced by Indiana's Principal Investigator, Daniel Reed, and will provide answers to the following questions:

- What has been accomplished in each area during this development cycle?
- How do these accomplishments relate to the project's objectives and timeline?
- How does project staff rate the quality of work that was done to produce these accomplishments?
- What problems or difficulties were encountered in each area? What strategies are being used to address these problems?
- Are there issues to be raised, or potential problem areas flagged, for the next development cycle?

The project timeline will be reviewed at the first meeting of ATM and AWM staff and will serve as a baseline from which to assess progress for the purpose of these evaluations. We do not necessarily expect ATM and AWM to follow identical plans or tackle every objective at exactly the same time. Our evaluation process will take this into account while, at the same time, identifying areas that require a shared timeline. These interim evaluations will be circulated to all project staff, including members of the Advisory Board. A general version of each evaluation will be posted on the project website so that interested outside parties and stakeholders can follow our development.

The Advisory Board for the *Sound Directions* project is another tool by which we will conduct an in-process evaluation. The Board will provide feedback on our workflows and preservation plans at the onset of the grant. When convened for the second time, the Board will evaluate the work that has been accomplished to date while providing further feedback for the duration of the project.

Finally, *Sound Directions* will make use of a summative method of evaluation in producing a final report that examines in detail the objectives, methods, results and impact of the project. This report will also point to the future, outlining further work that is needed to build on our discoveries and accomplishments.

VI. Work Plan

The *Sound Directions* project will unfold over eighteen months through a number of discrete stages with some overlap. The initial two months of the project will focus on preparation and consultation, including the first meeting of the Advisory Board. This stage will also feature collaborative discussion on the nature

of programming necessary to the project: can the ATM make use of software already developed at AWM? What are the programming needs at each institution and how should they be addressed?

During the second stage, hands-on research and development at both institutions will begin. Stage 3 brings the reconvening of the Advisory Board to add additional feedback and the incorporation of this information into the project. Dissemination of project results plus final evaluation and reporting take place in the last stage.

The following work plan highlights the different needs of our institutions that will model different workflows and adaptations. For instance, development of a database may be necessary for one while the other may have database functions in place. Because AWM has already invested substantially in developing metadata capture tools and managed digital storage for audio, hands-on engineering will start earlier than at Indiana, where the project programmer will begin several months ahead of the engineer.

Stage 1: Preparation, Consultation, Planning and Programming
Month 1 (February, 2005):
- Advertise programmer/database specialist positions at Indiana;
- Advertise the audio engineer position at AWM;
- Prepare documents for the Advisory Board that detail workflows and preservation systems at each institution;
- Organize Advisory Board meeting.

Month 2 (March, 2005):
- First meeting of the Advisory Board; first meeting of Indiana and Harvard project staff;
- Planning session for software development with project staff: assess the requirements of a preservation exchange package and determine where there are shortfalls at each institution and plan programming accordingly;
- Specification of metadata capture tools and workflow database requirements; assess necessary modification of existing tools/development of new tools for metadata;
- Advertise the programmer position at AWM.

Stage 2: Hands-On Research and Development
Months 3-5 (April-June, 2005): First interim evaluation due at the end of month 3.
- Programmers start—programming and database development begins;
- Project assistant starts at AWM; audio engineer starts at AWM;
- Initial development of project website;
- Reconfigure and upgrade studios at ATM;
- Advertise engineer position at ATM; advertise project assistant position at ATM;

Months 6- 8 (July- September, 2005): Second interim evaluation due at the end of month 6.
- Audio engineer and project assistant start at Indiana;
- Development/adoption of metadata tools and workflow database completed at Indiana;
- Development of ingestion tools begins at Indiana;

Months 9-11 (October-December, 2005): Third interim evaluation due at the end of month 9.
- Research and development continues in all areas; digitization continues;
- Development of ingestion tools completed at Indiana;
- Second meeting of Indiana and Harvard project staff;
- Research, development and programming on interchange of preservation packages begins;

Stage 3: Feedback on R&D and Mid-Project Corrections
Month 12 (January, 2006): Fourth interim evaluation due at the end of month 12.
- Re-convene Advisory Board;

Months 13-17 (February-June, 2006): Fifth interim evaluation due at the end of month 15;
- Digitization continues, additional Advisory Board feedback is incorporated;
- Refinements to metadata capture and ingestion tools;
- Additional development on public areas of project website;
- Digitization is complete at the end of month 17, project engineer positions end.

Stage 4: Reporting, Evaluation and Dissemination
Month 18 (July, 2006): Note that this stage will extend beyond Month 18 for Indiana and Harvard project staff who will be working on dissemination and wrapping up the project.
- Dissemination of project results; final evaluation and reporting.

VII. Staff

A. Indiana University
1. Co-Principal Investigator. Daniel B. Reed, Director, ATM .10 FTE for 18 months.

Ethnomusicologist Reed has experience on several digital archiving projects, including the NEH-funded CD-ROM *Music and Culture of West Africa: The Straus Expedition* (Indiana University Press 2002), the IMLS-funded Hoagy Carmichael Digitization Project, the LSTA-funded Starr-Gennett Digital Archive, and the Mellon-funded Ethnomusicological Video for Instruction and Analysis Digital Archive (EVIADA). Reed also served on the advisory committee for the Council on Library and Information Resources which created and evaluated the results of a survey tool to assess the state of audio collections in academic libraries around the U.S.

2. Project Manager. Mike Casey, Coordinator of Recording Services, ATM. .30 FTE for 18 months.

Mike Casey will serve as project manager for the Indiana side of the grant, guiding day-to-day work at the ATM while monitoring progress at the Digital Library Program. With the project engineer, he will be responsible for Indiana's research and development related to digitization. He will train and supervise both the project engineer and project assistant and will assume overall responsibility for quality control of the preservation work that is completed. Mike Casey has training and experience as both an audio engineer and a sound archivist. At the ATM he manages or performs all access/preservation transfer and restoration work for its 110,000 audio recordings dating from the 1890's to the present. From 1987 to 1993 he headed the Southern Folklife Collection in the Manuscripts Department at the University of North Carolina as the Department's Sound and Image Librarian. From 1994 to 1998 he worked part-time at UNC as a preservation transfer engineer while pursuing a music career. At UNC he completed several successful analog preservation projects, gaining extensive experience in the physical restoration and playback of deteriorating discs and tapes. He has learned directly from archival audio engineers such as Seth Winner and Tom Owen and taken audio engineering courses at various places including IU.

3. Digital library Program Project Coordinator. Jon Dunn, Associate Director for Technology, Digital Library Program, and Senior Technology Advisor, Indiana University Libraries. .10 FTE for 18 months.

Jon Dunn will serve as project coordinator for the work being carried out in the Indiana University Digital Library Program and will supervise the Programmer / Database Specialist. With the Programmer / Database Specialist and in cooperation with the Project Manager, he will be responsible for Indiana's

research and development related to preservation packages and long-term storage, and for the integration of support for the technical best practices and standards choices developed in this project into Indiana University's digital repository architecture. Jon Dunn has worked in digital music libraries for the past ten years, initially as technical director of the Indiana University Music Library's Variations digital audio library system, and more recently in IU's Digital Library Program, where he oversees technical work, including software development, system administration, and standards adoption, and is leading the development of an integrated digital library infrastructure for the university. He also serves as Executive Investigator and Lead Technical Investigator for IU's NSF-funded Variations2 digital music library research and development project and as Senior Technology Advisor to the Dean of University Libraries.

4. ATM Librarian. Suzanne Mudge. .05 FTE for 18 months.

Mudge will consult as needed on all issues of MARC cataloging, authority work, and archival access standards. With 10 years of experience as the ATM librarian and cataloger, Mudge is an active participant in Name Authority Cooperative Program (NACO) and works extensively with other librarians and scholars to develop descriptive cataloging standards and subject access for complex field collections. She has served as author, manager, cataloger, and consultant in numerous successful Archives grant projects, including ones funded by NEH, LSTA, and IMLS. She is currently a consultant and cataloger for the EVIADA project, leading its work on controlled vocabulary for effective searching.

B. Harvard University
1. Co-Principal Investigator. Virginia Danielson, Richard F. French Librarian and Curator of the Archive of World Music. .05 FTE for 18 months

An ethnomusicologist by training, Danielson has overseen the development of a state-of-the-art audio preservation studio at Harvard and launched the production of research-intensive multi-media documents constructed to international standards. She has served on the Board of the Association for Recorded Sound Collections and is an active member of the International Association of Sound Archives. She has authored a number of reports and articles dealing with digital access to multi-media collections, notably "Stating the Obvious: Lessons Learned Attempting Access to Archival Audio Recordings" the lead article in the collection *Folk Heritage Collections in Crisis* (Washington D.C.: Council on Library and Information Resources, 2001). As director of the Loeb Music Library at Harvard, she is responsible for a budget of $1.5 million and supervises 22 staff members. From 1988-1992 she administered a succession of federal grants to the Music Library for the purpose of retrospective conversion of music resources.

2. Lead engineer. David Ackerman, Audio Preservation Engineer. .05 FTE for 18 months.

Ackerman will guide the technological aspects of this project at Harvard. Working with Mike Casey and the Advisory Board, he will frame the questions, plan and supervise the work, and create the reports for the project. Ackerman is internationally known for his work on audio preservation metadata. At Harvard he designed and installed a state-of-the-art audio preservation studio; he performs or oversees all audio preservation projects in the Harvard College Library and consults on many others. He led the project "Music from the Archive" which contributed substantially to the development of standards for audio metadata, and created software for the collection of metadata and for the deposit of audio files in the Digital Repository as well as an internal SIP for the deposit of archival audio in the Repository. He has extensive experience reformatting virtually all forms of audio tape and discs. An active member of the Audio Engineering Society, Ackerman chairs the Working Group on Digital Library and Archive Systems, Task Group SC-03-06-A Metadata Harmonization.

3. Archivist and Intellectual Access Specialist. Sarah Adams, Keeper of the Isham Memorial Library. .10 FTE for 18 months.

Adams will oversee the intellectual access to the products of Harvard's project via EAD finding aids and catalog records. A scholar and specialist in rare musical resources, Adams adapted the EAD protocols to suit collections of audio recordings. She served in this capacity in the management of "Music from the Archives," where she worked primarily with ethnographic collections. Her work was presented as a paper at the International Association of Music Libraries in Tallinn, Estonia in 2003, entitled "EAD for 'International' Music: Digital Access to Archival Collections."

4. Constance Mayer, Public Services Librarian. .05 FTE for 18 months

Constance Mayer has combined an interest in digital libraries with leadership in public services for over ten years. As Head of User Services at Indiana University, she participated in the development of the Variations digital audio library system where her contributions focused on user interface design, communication with users, structural metadata for music, and workflow management. In her current position as Public Services Librarian at the Loeb Music Library at Harvard, her responsibilities include leading the organization of digital music resources on the web and managing a pilot project to digitize 30 rare and unique scores using systems and services developed by the Harvard Library Digital Initiative.

C. New Positions

1. Programmer / Database Specialist. Two positions: .5 FTE for 16 months at Indiana and .5 FTE for 15 months at Harvard.

The Programmer / Database Specialist will be responsible for database and software design, implementation and documentation. This work will include developing a database and software utilities to assist in managing efficient production workflows; developing and adapting software tools for metadata capture, creation of METS SIPs (submission information packages) and DIPs (dissemination information packages); and at IU the ingestion of digital content into the Fedora-based digital repository. The Indiana programmer will also be responsible for technical development in support of the project web site.

2. Project Audio Engineer. Two positions: 1 FTE for 12 months at Indiana and 15 months at Harvard.

Working under the supervision of the Coordinator of Recording Services at Indiana and the lead engineer at Harvard, the Project Audio Engineers will be responsible for research and development related to the digitization of archival recordings as outlined in the grant narrative. The engineers will also have primary responsibility for digitizing recordings carried on magnetic tape as well as lacquer and aluminum discs. This will include some work on physical restoration of project items (cleaning, baking, etc.) along with sonic restoration as appropriate. These positions will be responsible for extensively documenting both the source recordings and the transfer process utilizing emerging AES metadata standards.

3. Project Assistant. Two positions: 1 FTE for 12 months at Indiana .29 FTE for 15 months at Harvard.

The Project Assistants will be responsible for preparing collection documentation for digitization while managing workflow into and out of the preservation studio. At Indiana, the assistant will also mark divisions of content in files and collect appropriate metadata, then construct or update ATM indexes. Other duties at the ATM include checking derivatives, adding announcements to files, checking technical metadata, producing physical copies, and interfacing with the ATM Archivist and Librarian.

D. Advisory Board.
The Advisory Board will be convened twice during the grant period to provide feedback and outside perspectives as discussed in the Methodology section above. Members of the Board include Peter Alyea (Digital Conversion Specialist, M/B/RS, Library of Congress); Chris Lacinak (Director of Operations, VidiPax), George Massenburg (President, GML, LLC); Adrian Consentini (Chief Audio Engineer/Preservation Specialist, VidiPax), and Clifford Lynch (Executive Director, Coalition for Networked Information). A digital library generalist working on LC's Culpepper project will be invited to join the Board after the grant's competitive process is complete.

VIII: Dissemination

We hope to encourage utilization--the actual use of our results, discoveries and findings-- by disseminating widely with the plan outlined below. To achieve this goal we have designed a multi-level communication strategy and defined the characteristics of the content that we will disseminate.

A. Communication
Our working assumption is that one-way communication based solely on written information is important, but limited, in its effectiveness. Therefore, we have incorporated a variety of communication methods in our strategy including several that contain the possibility of two-way communication.

1. Project Website

The project website will be launched during the first stage of the grant and will, in effect, begin dissemination in the early part of the project. The website provides a tool for in-process dissemination and will include versions of our interim evaluations, allowing interested outside parties and stakeholders to follow the development of the project. The website will also host the final grant report and the edited booklet we will produce (see below). Finally, we will establish a direct feedback channel for potential users with a page that encourages questions and suggestions via email to various project staff.

2. Professional Organizations

We will formally notify various professional organizations when the project website is established and results are posted. We will also propose panels and talks to enable direct communication with potential users of our results. We will communicate about *Sound Directions* with the following organizations:

- Audio Engineering Society Working Group on Digital Library and Archive Systems, Task Group SC-03-06-A Metadata Harmonization that is working on metadata standards and is chaired by Harvard's David Ackerman. This group will be consulted early in the project as part of both the feedback and early dissemination processes;
- Association for Recorded Sound Collections, communicating through their listserv and proposing a panel discussion at the annual conference on digital preservation undertaken by institutional archives. Indiana's Mike Casey is co-chair of the ARSC Technical Committee;
- Society of American Archivists listserv along with a proposed panel for their annual conference;
- American Library Association through a proposed paper at their annual conference;
- International Association of Sound Archives through their listserv;
- Association of Moving Image Archivists and AV Media Matters listservs;
- Digital Library Federation through their listserv and twice-yearly forums;
- METS implementation registry at the Library of Congress;
- METS profile registry at the LC Network Development and MARC Standards Office;
- Joint Conference on Digital Libraries through their annual conference;

- Society for Ethnomusicology through their listserv and annual conference; Reed and Danielson are members of the SEM Archiving Section;
- American Folklore Society

3. Advisory Board

The *Sound Directions* Advisory Board, composed of potential users of our results, will help insure that the information we disseminate is relevant to the field. The Board will also act as an early dissemination tool by virtue of both their status as potential users and the fact that they also naturally communicate with many others in the field.

4. Publication

In addition to the final grant report, we will produce a booklet developed specifically for institutions holding collections of audio field recordings. Our working title for this booklet is *Sound Directions: A Blueprint for the Digitization and Long-Term Preservation of Ethnographic Audio Collections.* Even though this document will focus on ethnographic materials, we believe it will have even wider applicability since our research will explore many issues common to all audio preservation projects in the digital domain. We are currently talking with CLIR about their potential interest in helping us disseminate this booklet. Although the organization does not normally publish reports that they do not author, they have agreed to consider providing a link from their website to our publication.

B. Content
To reach a variety of potential users, our dissemination system will provide several types of information at different levels. For example:

- The booklet will include an executive summary that will frame our results in terms of the big picture. This section will be geared towards collection managers;
- The core of the booklet will consist of an in-depth discussion of the issues we researched and the results we obtained. It will provide significant detail so that the user can learn enough about both the basic principles and specific practices we used to pursue implementation in their own setting;
- The on-line version of the booklet will contain links to resources that might help implement the procedures and strategies presented;
- The booklet will reflect several dissemination channels including proactive--issues that users identified as important that we uncovered while developing this project or from the Advisory Board—and reactive, that is, issues that users may not have known were important that we uncovered during our research.

The final grant report will include all of the detailed information, including budgetary data, needed to provide an in-depth report on the objectives, methods, results and impact of the project to NEH. This report will also point to the future, outlining further work that is needed to build on our discoveries and accomplishments. The booklet, however, will be designed with institutional archives in mind. Along with a detailed exploration of project issues and results, it will include background information on the audio formats we preserved, assessing their strengths and weaknesses. This publication will be the first to provide a project-based examination of critical areas in the digital preservation chain that is both detailed and comprehensive. It will represent a significant contribution to the development of standards and best practices for the preservation of audio recordings in the digital domain.

Notes:

[1] Council on Library and Information Resources. "The State of Audio Collections in Academic Libraries." Proposal, 2002. Page 1.

[2] Ibid.

[3] International Association of Sound Archives. Technical Committee Papers. "The Safeguarding of the Audio Heritage: Ethics, Principles and Preservation Strategy." Version 2, September 2001. Standards, Recommended Practices and Strategies IASA-TC 03. At http://www.iasa-web.org/iasa0013.htm

[4] A very useful discussion by Carl Fleischhauer of the LC prototyping project and the development of Culpepper is in the proceedings from Sound Savings: Preserving Audio Collections, a symposium held at the University of Texas-Austin in 2003. The document is located at:
http://www.arl.org/preserv/sound_savings_proceedings/fleischhauer.html

[5] Cohen, Elizabeth. "Preservation of Audio." *Folk Heritage Collections in Crisis*. Washington, D.C. Council on Library and Information Resources, 2001. Page 26.

[6] We examined best/recommended practices documents from the following: the International Association of Sound Archives, Library of Congress, Colorado Digitization Program, Research Libraries Group, Arts and Humanities Data Service, National Initiative for a Networked Cultural Heritage, MATRIX at Michigan State University, University of Texas, University of Kansas, University of Wisconsin, Cornell University (Macaulay Library of Natural Sounds) and the book titled "Creating Digital Audio Resources: A Guide to Good Practice." We also examined digitizing projects at a number of institutions including: McGill University, University of Missouri—Kansas City, University of California—Santa Barbara, National Library of Norway, National Library of Australia, Bulgarian Academy of Science, University of Texas—Austin, and the PARADISEC project (Pacific and Regional Archive for Digital Sources in Endangered Culture) and PRESTO Project.

[7] As our proposal was in production we received a draft of the IASA best practices document, IASA-TC04 Guidelines on the Production and Preservation of Digital Audio Objects, due to be published in August, 2004. This is the most useful best practices document we have seen--it will certainly become a key digital audio preservation reference. It provides a wealth of detail on the playback of source recordings, an area that we are not exploring. It also provides detailed recommendations in several areas that overlap with *Sound Directions*. However, our project is designed to provide very specific and detailed information in these areas drawn from project work rather than the guiding principles found in TC04. In areas like the collection of metadata, for example, we will develop procedures for using specific metadata gathering tools while also exploring questions of who collects metadata, at what stage and the resources and expertise that are needed. In effect, we will be putting into action the IASA principles, using them to produce detailed practices and procedures—nuts and bolts information—that will aid other institutions as they learn how to build their systems in these areas. Other areas that we will explore in our project, such as the construction and interchange of preservation packages and quality control for example, are not covered in the IASA document.

[8] The document "Digitizing the World's Largest Collection of Natural Sounds: Key Factors to Consider when Transferring Analog-Based Audio Materials to Digital Formats" from Cornell University explores quality control as it relates to physical media but does not address the digitizing work itself. Library of Congress documents also address quality control to some extent.

[9] The Safeguarding of the Audio Heritage: Ethics, Principles and Preservation Strategy, Version 2, September 2001. Standards, Recommended Practices and Strategies. IASA Technical Committee, IASA-TC 03. Located at: http://www.iasa-web.org/iasa0013.htm

[10] Note that we received an advance draft of the forthcoming IASA best practices document TC04 while this proposal was in production. We will study this document and incorporate its recommendations into our workflows, using a real world project to build on the work of the IASA technical committee.