

Thinking Outside the Archive: Collaboration and the EVIA Digital Archive Project

Alan R. Burdette, *Indiana University*

Based on paper presented at the IASA International Conference, Sydney, Australia, September, 2008

Introduction: The necessity of collaboration to an archive

The landscape for sound and media archives has changed dramatically in the last decade. Digital audio preservation has moved from being a question, to the creation of best practices, to being an imperative. With several media formats from several different generations on the brink of deterioration or obsolescence in the next decade, digital preservation is a necessity that we cannot ignore. Indeed, even the most recent born-digital “media-free” recordings are also in grave danger, not so much because of format obsolescence but because of poor management by individual collectors. Video is in even greater danger than audio holdings due to the greater complexity of the playback technology and the higher degree of format obsolescence. No less than preservation concerns are the pressures for online access to collections and the accompanying challenges to do so legally, ethically, and in ways that are easy and attractive to users. Unfortunately, most media archives are under funded, understaffed, and dependent on grant funding for any kind of preservation or access initiatives. This is an exciting but turbulent time for media archivists as they retool for new technologies and new partnerships.

Collaboration is not new to archives. Filmmakers, scholars, publishers, record producers, commercial vendors, and museums have all been collaborating with media archives for decades. However, the new paradigms of digital preservation and access have required collaborations on a scale and at a depth that is unprecedented. Indeed, preservation and access requires partnerships that stretch the boundaries of the conventional sound archive in ways that challenge some of the very notions of who we are.

The EVIA Digital Archive Project is an excellent example of broad collaborations that have been successful in addressing several key issues for the Archives of Traditional Music and for the field of ethnomusicology and related ethnographic disciplines. Working with technologists, software developers, scholars, publishers, university administrators, lawyers, librarians, other institutions, vendors, editors, and other archives; the project is addressing the most fundamental concerns of archives as well as reaching towards new modes of archival action.

Overview of the Project

The EVIA Digital Archive Project is a collaborative effort to establish a digital archive of ethnographic field video recordings and an infrastructure of tools and systems supporting scholars in the ethnographic disciplines. With a special focus on the fields of ethnomusicology, folklore, anthropology, and dance ethnology, the project has developed a set of tools for use by scholars, instructors, librarians, and archivists. In 2001, the project was funded by the Andrew W. Mellon Foundation, Indiana University, and the University of Michigan, and built through the collaboration of ethnomusicologists, archivists, librarians, technologists, and legal experts. The primary goal of the EVIA Project has been to preserve ethnographic field video created by scholars as part of their research. Its secondary goal has been to make those materials available online in conjunction with descriptive annotations. Project staff and contributors have created a support system and a suite of software tools for video annotation, online collection searching, controlled vocabulary and thesaurus maintenance, peer review,



and technical metadata collection. The project will reach the end of its funded implementation phase in June of 2009 and has begun to provide online access to the finished collections in the spring of 2009.

The EVIA project is unusual in that it was allied with the Archives of Traditional Music, but it was not primarily about preservation and access to existing ATM holdings. The impetus for the project came from its co-principal investigators Ruth Stone and Lester Monts and their recognition that they and many other scholars had significant video collections in their personal possession that had not been deposited into an institutional archive. Part of the project's mission has been to draw those collections into research archives.

Ruth Stone at Indiana University and Lester Monts at the University of Michigan are scholars of Liberian music and both were concerned about their own deteriorating collections of video recordings. They gained the attention of the Andrew W. Mellon Foundation which was also looking for a promising video project to support. With initial planning phase funding, in 2001-2002 they brought together a diverse group of experts for three intense planning sessions over the course of nine months. The results of these planning sessions established the framework for the project, its mission, and the collaborative principles that would be necessary for success. The planning sessions were tremendously productive but also challenging. Many different agendas were present and not everyone was in agreement.

Through the course of the year-long planning phase and then through several years of development, the EVIA Project has created an ambitious agenda to address a broad range of scholarly needs in ethnomusicology and related fields. These needs include infrastructure for digital video preservation and delivery, new models for scholarly communication and publication, and solutions for the technical and legal challenges required by the creation of an online publishing system for archival materials.

Preservation

The preservation of video recordings is at the core of the EVIA project mission. Video began to be adopted by scholars in the mid-1970s and widely employed by the late 1980s, because it offered an inexpensive way for researchers to capture a fuller range of expressive culture than still photography or audio recordings allowed. However, the deterioration of video signals on various carriers is dramatically fast compared to audio, as is the obsolescence of playback equipment. For example, because the Hi-8 format enjoyed less than ten years as a popular field recording format, and because it was not used in the professional broadcast industry, playback equipment has disappeared quickly and is now difficult to find. We have also seen a greater obsolescence of video recording formats compared to audio. Between 1990 and 2008, the formats of VHS, VHS-C, 8mm, Hi-8, and MiniDV were all popular with scholars as field recording. Today, MiniDV is the only one of these formats that is not considered obsolete as a field recording format, but it, too, will soon be left behind as the industry moves quickly towards tapeless recording solutions.

While audio archivists have come to broad consensus about digital audio preservation as described and proscribed in *IASA TC-04*, and in *Sound Directions, Best Practices for Audio Preservation*, the same has not yet been achieved for digital video preservation. Digital preservation of video within the EVIA Project and many large media archives has proceeded in the absence of broadly accepted best practices, but has used the model of audio preservation as well as a careful assessment of the best way forward in the absence of such guidelines. We have been careful to adhere to basic digital preservation principles in our formulation of a solution to long-term digital video preservation and have consulted with other digital

video preservation efforts around the world. Unfortunately, the disintegration of video recordings requires that we act now to make preservation transfers rather than wait until standards and best practices are adopted. One of the key areas of partnership in the EVIA project has been to utilize the facilities and the expertise of the Duderstadt Media Center at the University of Michigan for videotape transfers. All but the oldest and most unusual formats are digitized there and then sent to Indiana University for quality assurance. Formats such as 1/2" open-reel or U-matic have been sent to commercial vendors.

Access

The archival holdings of the EVIA Project are accessible online but are restricted to educational uses. Rather than making these holdings widely available, they are available through selected academic institutions, groups, or by individual permission. Most of the recordings in the EVIA Project were recorded under the condition that they only be used for research and educational purposes. We take the stewardship of these recordings seriously and so we carefully control access and require that users agree to a statement about how they may utilize the archive materials. The EVIA Digital Archive Project has endeavored from the very beginning to improve the ability of scholars to document their video recordings and to enable library-based searching of its content. Librarians have been a key part of the project design and we have incorporated standard MARC record cataloging as well as controlled vocabularies into the project implementation.

Scholarly Annotation and Publishing

At present, the EVIA Digital Archive consists of a group of video collections that have been selected by an editorial committee for inclusion in the project. The archive is designed to be a long-term preservation repository for these recordings as well as a unique peer-reviewed scholarly publication. In addition to the approximately 10 hours of video in each collection, scholars have worked extensively with their recordings to describe and analyze what they have documented. The content of the archive represents the culmination of preservation, annotation, and editorial work.

Annotation of video emerged from the beginning of our planning as a way to improve the quality of documentation the archive acquires with video recordings and the project has built tools to maximize the ability of a scholar to describe their video. This process is intended to help the scholar with their research and ultimately create a much richer resource for others. After working with our first group of scholars, we realized that we were also creating a new kind of digital publication and have since built additional software tools and administrative support for the peer review and publication process. Scholarly annotation of ten hours of video and the accompanying citations and glossary entries have regularly reached the equivalent of 80 pages of single-spaced text. By bringing to bear the conventional mechanisms of peer-reviewed publishing to video annotation, we are placing greater scholarly value on those annotations, and at the same time, transferring some of the typical academic rewards for peer-reviewed publishing to archival documentation.

Project Status

We presently have forty-five collections in different stages of completion as part of the archive. At the time of this writing, eight collections are ready for publication and will be made available as part of the first round of access to the EVIA Digital Archive. The small percentage of collections that are completed reflect the developmental nature of the project, and it is worth noting that projects that started in the more recent summer institutes have

reached completion more quickly than those of the earlier summer institutes. This success rate has to do with greater reliability of the software, better training, clearer models, and more efficient administration of the summer institutes and the ensuing follow-up. Because the annotation process is extensive and time-consuming, we have held two-week long summer institutes during which we bring a group of a dozen scholars together to work on their video collection. These summer institutes have proven to be very effective for training scholars to use the tools and to enable them to get most of their collection prepared. However, no scholar has completed all of the work at the summer institute and they continue to work on their collection after the institute is complete. The summer institutes have not only proven to be an important means for scholars to focus intensely on the task of annotation, but they have also been excellent opportunities for scholarly dialog between a diverse group that might not come together otherwise.

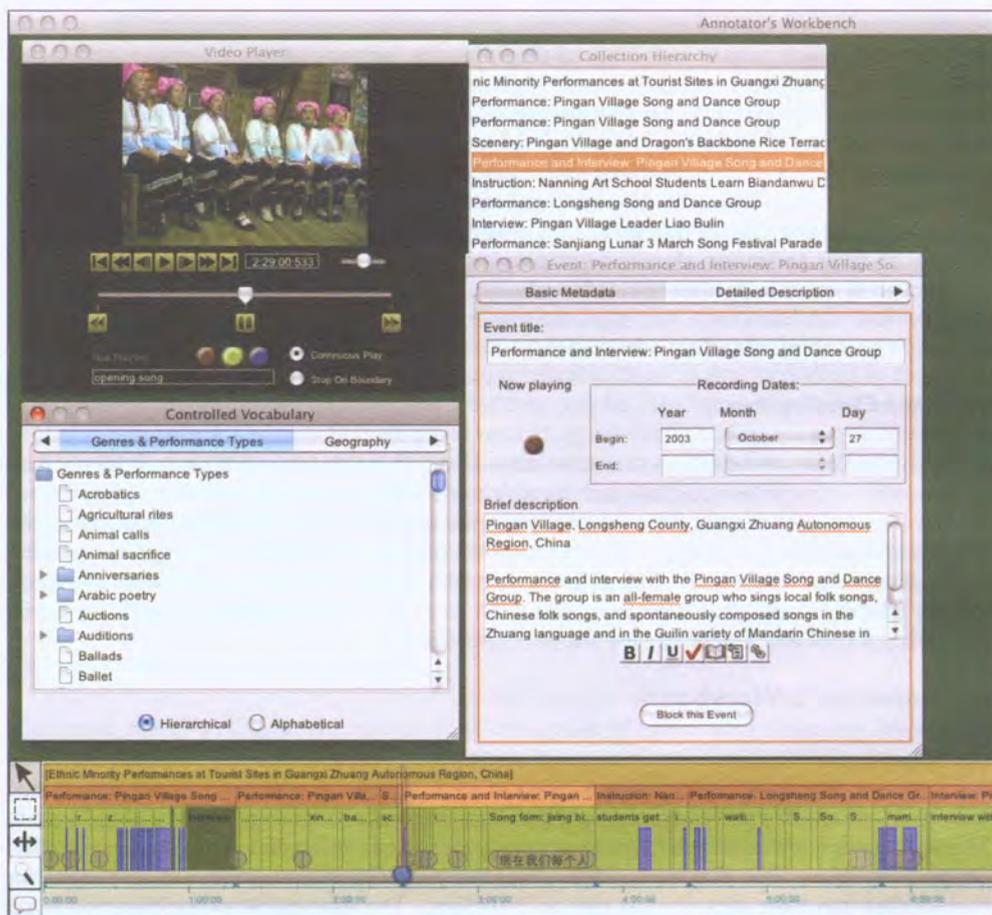
Software Development

Software development has been the most extensive and most expensive part of the project. At the time we began production, satisfactory tools did not exist or were too proprietary for our archival concerns. Software development, which will ultimately all be available through an open source license, has been focused on the production of a video annotation application and an online search and browse tool. Additional applications have also been built to address needs that have emerged such as a controlled vocabulary and thesaurus maintenance tool and a peer review support application.

The Annotator's Workbench

The Annotator's Workbench enables users to create a collection from a set of existing video files, segment that collection at several levels, create annotations for those segments, assign controlled vocabulary terms to segments, and to control access to parts of the collection.

The Annotator's Workbench is installed locally and reads video files that are also installed locally. The local video files are 1 Mbps (Megabit per second) transcodes of preservation masters that may have data rates as high as 270 Mbps. The Annotator's Workbench creates an output file that contains all of the metadata generated by the user in a METS-compliant .xml file. Once a project is complete, this file is archived with the preservation masters and is uploaded into our FEDORA repository for preservation and online delivery. Key features of the Annotator's Workbench include the ability to create a collection out of multiple files, segment that collection into intellectually meaningful units and layers, annotate those segments and layers with descriptive information and controlled vocabulary terms, create textual transcriptions and translations that are assigned to segments, and limit public access to segments within the collection if desired.



Caption: Screen capture showing several features of the Annotator's Workbench software developed by the EVIA Digital Archive Project.

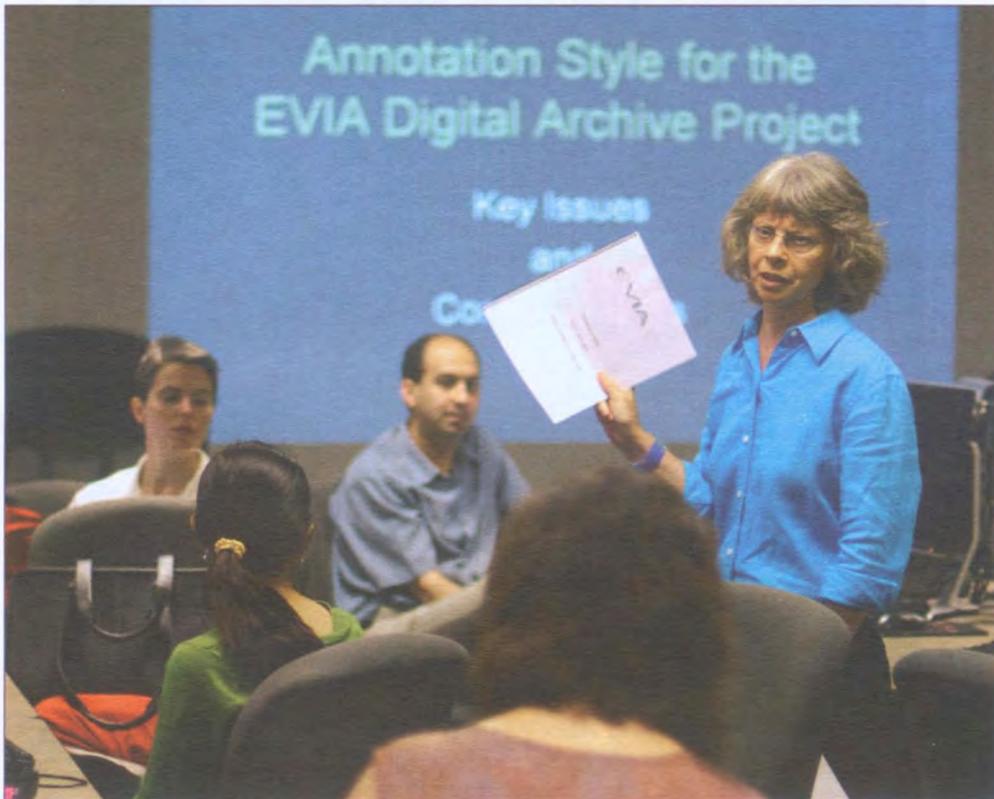
Online Search and Browse

Video collections created by the EVIA Project are accessed using our Online Search and Browse tool. This multi-functioned application allows users to search, browse, and access the video collections and multilayered metadata created by the scholar in the Annotator's Workbench. Full-text searching is available as well as browsing through assigned controlled vocabulary terms. Extensive playback controls, full-screen playback, and playlist creation are also available to the user. The multiple layers of annotation and transcriptions are all accessible through the interface.

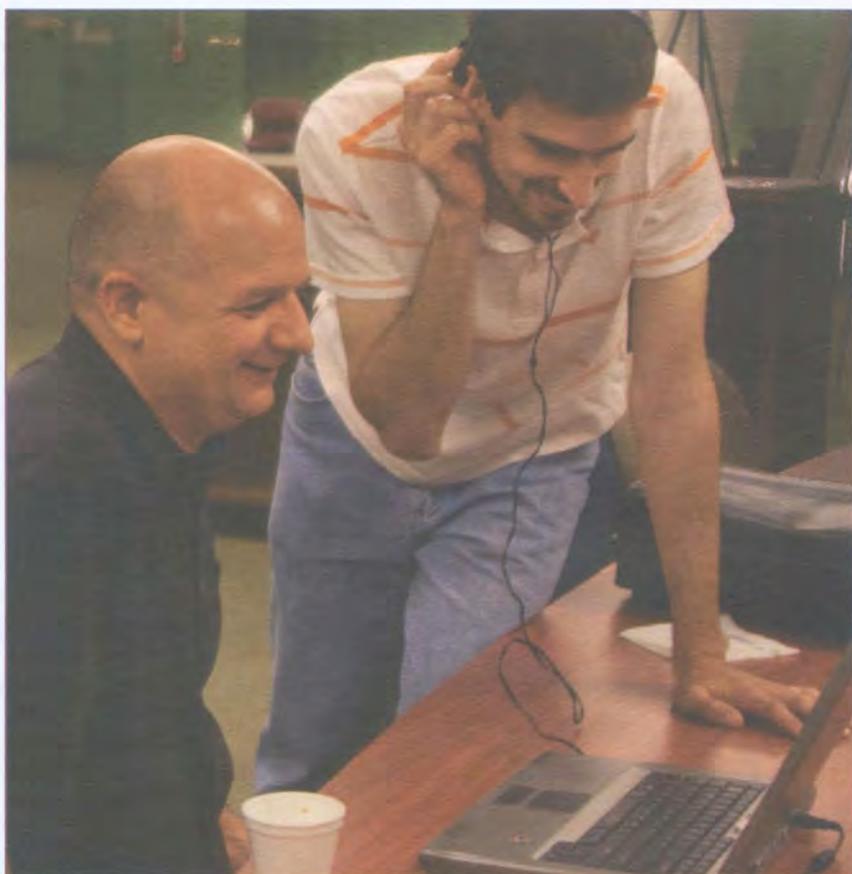
Controlled Vocabulary and Thesaurus Maintenance Tool

The Controlled Vocabulary and Thesaurus Maintenance Tool is designed specifically for librarians to create and manage controlled vocabularies and thesauri that will be used within the Annotator's Workbench application. The EVIA Project has designed its own controlled vocabulary categories, but the Annotator's Workbench has a modular design that allows

different controlled vocabulary schema to be used for different kinds of projects. This tool not only allows a librarian to maintain a list of controlled vocabulary terms and their authority references, but also enables the creation of different vocabulary sets and schema. During summer institutes and the period that follows, a librarian is in direct contact with the annotating scholars and negotiating the addition of controlled vocabulary terms to the existing lists.



Caption: EVIA Digital Archive Project assistant Colleen Haas gives a presentation on issues of annotation writing style during the 2008 summer institute at Indiana University.



Caption: Frank Gunderson and Alex Perullo compare video and annotations during the 2006 EVIA Digital Archive Project summer institute. Both Gunderson and Perullo are ethnomusicologists who have conducted research in Tanzania.

Sustainability

The EVIA Digital Archive Project is being sustained through collaboration with the Institute for Digital Arts and Humanities, the Digital Library Program, the Archives of Traditional Music, and the University Information Technology Services at Indiana University. In this way we have a base infrastructure from which we are applying for further grants to develop content, and are also using the EVIA architecture as a base platform from which to develop further functionality for scholarly media. For these sustainability efforts we are exploring collaborations with other institutions and projects.

From a preservation point of view, we are relying on the Archives of Traditional Music and the technology infrastructure at Indiana University for the long term viability of the files we are creating. As the Archives of Traditional Music generates digital preservation masters, we have become reliant on Indiana's Digital Library Program to maintain and migrate those files. At first glance it might seem as if the archive is ceding stewardship of those files to the Digital Library Program and the university because those files will no longer reside on

a server maintained by the Archives of Traditional Music. However, it is more accurate to view this situation as one in which the Archives and the Library have committed to long-term collaboration to manage these preservation masters, learning from each other about preservation and storage needs and requirements.

Collaboration

Within the EVIA Project, collaboration has occurred at multiple levels. More than just a case of two institutions pooling resources, we have created a collaborative framework that reaches out to many different units within our campuses as well as across many different institutions. A fundamental point is that a close collaboration with scholars has led the Archives of Traditional Music into greater technical collaborations as we work to support our mission principles. We haven't changed the mission of the archives in any radical way; rather we are better realizing our basic mission of preservation and access.

The EVIA project has been advanced during its development by institutional collaboration at several levels. At a basic level it has been a collaboration between Indiana University and the University of Michigan. Both institutions have provided significant resources in the way of personnel time, space, and expertise. The Mellon Foundation, of course, has also been a significant partner. They have not only invested significant funds into the project, but have also provided extremely important guidance to the project. The idea to hold a summer institute, for example, came from our Mellon program officer. The Mellon Foundation has also consistently pushed for solutions to long-term sustainability of the project. Another kind of institutional collaboration is emerging as we discuss with other archives ways in which we can share resources. In particular, archives that have good digital preservation copies, but lack an online access system are excellent partners with the EVIA project because collaboration is mutually beneficial to both.

Archives work with scholars all of the time, but primarily in one-way exchanges. Either archives serve scholars by providing them with materials they need for their research, or archives accept the collections of scholars for deposit. Scholars often describe their media items based on their own idiosyncratic systems and archives must translate these systems (if they indeed exist) into some archival standard. The depth of this kind of documentation is highly variable and often takes place long after the items were created. The EVIA project has endeavored to create tools and procedures for the documentation of video that supports the work of scholars during their research while simultaneously preparing them for archival deposit. We achieve this by using a METS/MODS compatible schema behind the various descriptive metadata areas that scholars utilize.

The EVIA project has been developed by multiple units across the campuses of both Indiana University and the University of Michigan. From a planning and development perspective, the project has depended upon the advice and regular feedback from scholars in the Ethnomusicology programs at both institutions. In the case of Indiana University it has involved nearly all of their faculty and several graduate students in the department. In this sense, the department has invested itself in the success and development of the project.

Indiana University's Digital Library Program has been key to keeping the project development aligned with library standards and the ever-developing cyber-infrastructure of the university. By insuring that project development uses the same technologies of other library projects they enable sharing of code and functionalities across several projects at Indiana University. This helps the software developed by EVIA take advantage of related software development for other projects, and EVIA technology can also be more readily integrated into other projects.

Without the infrastructure and technology expertise provided by Indiana's Digital Media Network Services (DMNS) and the University Information Technology Services (UITS), the EVIA project would not have been possible or at least affordable. UITS is the centralized Information Technology service for the entire campus and DMNS is a sub-unit within UITS that provides network and streaming media support. UITS provides the Mass Data Storage System (MDSS), a data tape robot with a current capacity of 2.2 petabytes. The EVIA project uses MDSS for long-term storage and it will be the backbone of the preservation repository service of the Digital Library Program that will be implemented by the end of 2009. We could not have seriously considered the EVIA project without these key resources and without the support of UITS. By the same token, the EVIA project has been an important part of broader arguments at Indiana University for high capacity digital storage for the humanities and EVIA has also been a strong part of the rationale for preservation repository services. As much as we have needed the infrastructure provided by our universities, we have in turn helped justify a variety of technology infrastructure agendas.

In 2008, the EVIA project became officially affiliated with Indiana's new Institute for Digital Arts and Humanities (IDAH). By doing so, administrators and software developers employed by IDAH will maintain the basic functions of the EVIA project with seeking out and developing new sources of funding and collaboration for growth of the EVIA project. As the same time, the EVIA project provided some of the framework and models for IDAH and many of the experiences gained while administering the EVIA project were critical to informing how IDAH needed to be managed. One of the persistent difficulties on the EVIA project and any other digital project that relies on software developers hired by grant funds is the fact that good programmers are hard to find and hard to keep on soft money. As the end of any grant term nears, it is very common for programmers to look for new employment in order to insure that they won't be caught without a job if grant funding ends. For the EVIA project and many others like it, this has meant that project administrators spent a great deal of time hiring and training new employees and it has resulted in a lot of down-time on the project. The EVIA project has helped build a new model within IDAH where a stable of programmers are guaranteed a regular contract, but their time will be assigned to various grant projects or internal projects as needs arise.

Although the EVIA project began with a group of ethnomusicologists, they recognized in the planning phase that ethnomusicology was not a discipline that was large enough to sustain this work by itself. Thus, the project has worked to involve scholars from other related ethnographic disciplines such as anthropology, folklore, ethnochoreology, and various area studies. The long-term effect will be to broaden the base of users and create an interdisciplinary community that will better support the further development of the project.

The EVIA Project has actively sought collaboration with other projects that have preservation and/or access needs for video collections. In some cases EVIA has provided a ready-made preservation and access system for a video collection. The Archive of Historical and Ethnographic Yiddish Memory (AHEYM) project is working with the Archives of Traditional Music and the EVIA project to provide these preservation and access services. Nearly 700 hours of video from AHEYM will be available through the EVIA online collections but it will also be available through a special AHEYM portal they will create based on the EVIA Project architecture. In a related case, the EVIA project will be collaborating with the team from the Cheyenne Sacred Language project to provide similar preservation and access services to their project which is using video to document an endangered language. The EVIA project has also been collaborating with the Central American and Mexican Video Archive (CAMVA) project for two years in a mutually beneficial arrangement that has supported additional software development for the EVIA project while providing annotation and cataloging



software technical support as well as a customized access interface for the collaborating institutions in the CAMVA Project. Several other collaborations are in discussion and development, and we see this as an important way to expand content and continue software development.

The EVIA project sees many possibilities in the area of collaboration with the subjects and communities documented in the video recordings. In one case a scholar brought one of her video subjects to a summer institute to discuss the recordings and translations. In other cases, scholars have been working in the field to annotate and discuss their descriptions with the people they are documenting. We are very excited about the possibilities for integrating the research, publication, and archiving process more closely together. This kind of collaboration facilitates better documentation and finer control of access and permissions and is ultimately better for the field and better for the communities we study. It also has the potential to remove more of the ambiguity regarding access that plagues many ethnographic collections.

Collaboration Benefits

At a fundamental level, collaboration requires the members of an organization to confront the different ideas, workflows, and priorities of another organization and its culture. In the case of a good collaborative partner, one can be pulled forward by this confrontation and re-evaluate one's own status quo. Even when there is disagreement, the evaluation that precedes it is extremely helpful in defining one's perspective and practice in contrast. When there is agreement, those strengths can be shared and advanced in a more confident way. Ultimately larger best practices can be built this way.

In the eight years since the project began, the university has created a significant amount of bureaucracy to address financial accountability relative to grant management. However, from the point of view of the products of grant funded digital projects in the humanities, it is fairly easy to think of projects we have seen where the results were minimal and never fully realized or publically distributed. Despite whatever reports and accounting the funding agency requires, single institution projects do not experience the kind of accountability that is naturally built into a collaborative project. Working together with another organization puts one's credibility on the line and the friendly competition is useful for pushing the goals of the project forward.

One of the unexpected benefits of collaboration within the EVIA Project has been between scholars themselves. Scholarly activity within the humanities still tends to be a very solitary endeavor and it is rare for scholars to engage in extended dialogs with their colleagues that result in tangible products. The tools and procedures of the EVIA project are the result of many years of back and forth dialog with scholars about fundamental issues of ethics, the uses of video, and digital publishing. The creation of a controlled vocabulary schema required an extended discussion among a group of scholars that caused them to analyze and think through their own assumptions about disciplinary categories. In many ways we have only scratched the surface of scholarly collaboration and anticipate that as the project broadens its disciplinary horizons, we will see more interdisciplinary questions and collaborations arise.

Archives often have difficulty justifying their existence relative to their costs and the amount of use they actually see. Through the EVIA project, the Archives of Traditional Music has helped make a case for special collections generally speaking and has moved the institution forward in very useful ways. By thinking broadly and consistently making the arguments for

how the development of preservation and access services serve the entire university, archives demonstrate their value to their larger institution. The preservation of video or audio follows the same principles regardless of the content, and while surface features of access may be different, the underlying technologies of delivery are fundamentally similar.

Collaboration Challenges

Good communication is essential to effective collaborations and it is where most collaborative difficulties lie. We have certainly struggled within the EVIA project to keep all of the stakeholders apprised of what was going on. With distance, face-to-face meetings become difficult, time-consuming, and expensive. Even with easy access to all manner of electronic communication, face-to-face meetings are extremely valuable to building effective collaboration and to keeping it alive. Projects or aspects of a project can run off the rails very easily if not attended to with regular communication.

Archives thrive on careful controls, but collaboration inevitably involves compromise and by extension some loss of control. If a collaboration is set up carefully, then a project can gain from a partner who is better at controlling that aspect of the project. However significant problems can arise when an archive cedes key aspects of its mission and its reputation to an organization that is not better at controlling its aspect of the project. This is a recipe for frustration on all sides of the collaboration and this is why collaborative partners must be chosen carefully.

One of the most important challenges to collaboration is the continual evaluation and demonstration of how a given collaboration is mutually beneficial to the parties involved. Collaboration breaks down when partners no longer see its benefit and do not invest the energy in its success. The partners need to strike a balance between selfishness and selflessness and be open-minded about the give and take that is necessary.

Conclusion

Archives are places that hold treasures of enduring value, and the preservation of these holdings and the collective access to them benefits many. Ultimately, archives will be judged by what benefits they are providing to the public, but many of us struggle with the fact that the users of ethnographic media archives are often dispersed and esoteric in their interests. We preserve and protect those things that only a few find valuable now, recognizing that they will be invaluable to some in the future. The more successful we are at finding allies in many corners, the more successful we will be in our more basic mission of preservation. At present, digital preservation demands collaboration for all but the largest archives. Beyond the technological and economic necessity for collaboration, however, is a further challenge to work with a broader base of scholars, publishers, librarians, technologists, and users so that we will be better prepared to sustain our collections into the future. In so doing, we make them more invaluable and more present in the lives of many. We can no longer afford to perpetuate the perception of archives as a container to put things in, but must see what we do as a kind of public action that is integral to the broader mission of other institutions and to the everyday work of our constituents.

References

Casey, Mike and Bruce Gordon. 2007. *Sound Directions: Best Practices for Audio Preservation*. Indiana University and Harvard University.

<http://www.dlib.indiana.edu/projects/sounddirections/bestpractices2007/>

International Association of Sound and Audiovisual Archives. Technical Committee. 2009. *IASA-TC04 guidelines on the production and preservation of digital audio objects: standards, recommended practices, and strategies*. Second Edition. Aarhus, Denmark.