

Current Practices in Digital Asset Management

version 0.9



The Internet2/CNI
Performance Archive & Retrieval
Working Group
October 2003

Current Practices in Digital Asset Management

Table of contents

Introduction.....	1
Background	1
Who Should Use this Document.....	1
Analog vs. Digital: Considerations for Digital Asset Management.....	2
Getting Started	4
General Resources and Requirements	4
Resources Specifically for Asset Management.....	4
Current Practices in Institutional Policies	5
Preservation.....	8
Intellectual Property	11
Rights and Licensing.....	12
Privacy.....	14
Fair Use.....	15
Metadata Standards, Schemas, Frameworks, and Formats	16
Some Commonly Used Metadata Schemes.....	17
Repository Frameworks and Structural Encoding Standards	18
Standards For Encoding Metadata for Resource Discovery.....	20
Content Standards.....	20
Too Many Standards?	21
Search and Retrieval	23
Digital Collection Software/Systems.....	24
Technology.....	24
Solution.....	25
Putting It All Together—Schematic System Configuration	27
Appendices.....	28

Introduction

Background

The mission of the Performance Archive & Retrieval Working Group is to propose standards and best practices for documenting, archiving, and retrieving the recordings of performances in the new digital environment. From the charter:

Advanced networks in combination with other developments in information technology, such as digital video/audio capture and synthesis, make it possible to document live theatre, musical compositions, dance and other performance in new ways. These developments also allow for the creation and documentation of new genres of performance such as collaborative distributed musical theater events. These materials will be important in the future scholarship of the arts and humanities. They will be an integral part of new theses and dissertations and also of scholarly publications. [Introduction 8]

As part of its output, the working group has drafted two guides, one for capturing live performance events and one for managing the digital assets created as a result of that capture.

The aim of this document is to provide a basic primer for people archiving and/or collecting digital recordings of live performances and performing arts events in research and/or educational environments. Specifically, we focus on faculty and artists within performing arts departments and organizations who find themselves in the position of creating digital archives of performing arts events. We expect that these individuals will be members of teams with this objective in mind; we intend to highlight the issues such teams will face and the areas where they will need to collaborate with others within their parent organizations. Although we provide a fair amount of detail, our goal is not necessarily to teach people all of the technical details involved; rather we hope to give a thorough overview of all of those details. We wish to be informative enough that, upon reading this guide, users feel that they know the right questions to ask fellow team members and other expert staff, and can start building their own expertise.

This document has been produced by the Performance Archive & Retrieval Working Group [Introduction 7*], jointly sponsored by Internet2 [Introduction 4] and the Coalition for Networked Information (CNI) [Introduction 1]. Internet2 is a consortium being led by over 200 universities working in partnership with industry and government to develop and deploy advanced network applications and technologies, accelerating the creation of tomorrow's Internet. It is recreating the partnership among academia, industry and government that fostered today's Internet in its infancy. CNI is an organization dedicated to supporting the transformative promise of networked information technology for the advancement of scholarly communication and the enrichment of intellectual productivity. Some 200 institutions representing higher education, publishing, network and telecommunications, information technology, and libraries and library organizations make up CNI's membership.

Who Should Use this Document

This document is written for faculty and other scholars of the arts concerned with how to best archive digital recordings of live performance in research and educational settings. It is expected that these faculty will find themselves in teams consisting of members with several areas of expertise including, but not limited to:

* This document's bibliography is organized according to its section headings. Within each section, references are arranged alphabetically. So, for example, the pointer "[Introduction 7]" refers to the seventh listing in the Introduction section of the bibliography.

- A) University departments, institutes, centers, etc. in several fields (such as other performing arts disciplines, anthropology and ethnology, ethnomusicology, folklore, etc.) which create and study performance and know the requirements and special challenges of those conducting research using non-textual, time-based source materials.
- B) Media centers and departments that understand the technical issues of digital capture and provide production and post-production environments that are ideally integrated (or at least compatible) with later content and collection management applications.
- C) Libraries and archives charged with preserving and organizing digital collections, as well as with providing access to the collections and tracking the legal and contractual obligations governing rights to and restrictions on such access.
- D) University information technology centers responsible for maintaining the hardware and software operating environments of such repositories (including technical backup and storage services), as well as for the database, application, and network environments necessary to implement each aspect of content management, from the capturing of raw streams to end-user playback of the final, archived objects.

We expect that such inter-disciplinary teams may also be interested in our guide for the capture of live performance events. Again, the purpose of this guide is not to teach faculty to become experts in preservation management, intellectual property law, metadata standards, repository architecture, and so on, but rather to give them the background necessary to be able to talk intelligently with experts about what they would like to accomplish.

Analog vs. Digital: Considerations for Digital Asset Management

This is a guide to current practices. Because the field of digitally-recorded performance is so new, we do not feel confident that there has been enough time for institutions to codify ‘best practice’. For example, there is no general consensus that digital formats have the qualities that would make them ideal formats for preservation/archival purposes (at least, not yet). Nevertheless, we forgo recommending analog formats (even popular ones) in favor of limiting our recommendations to digital formats. The reasoning behind this decision is hinted at in the quote from the Working Group’s charter that opens this document. While analog collections of live performance recordings have been—and remain—vital to our artistic heritage, many of the newest methods of creating and documenting performance are enabled by digital technology.

Many attitudes about the digital delivery of resources are based on a linear, passive, one-dimensional mode of thought shaped by commercial television. However, digital capture offers the opportunity for a richer set of options and products. Among these is the delivery of raw streams of output and/or multiple derivative resources created from the captured data. One of the exciting features of the digital domain is that a variety of outputs can realistically be produced in cost effective ways. For instance distributed multi-site performances and the use of content-based audio/video search and retrieval tools are both made possible by the availability of advanced networks to distribute digital performance collections. Multi-sensor capture can be delivered to users much more meaningfully in the digital environment.

Indeed, the steps of the digital media production process parallel those of the digital content lifecycle. In the digital media production process, raw feeds from capture go through editing and derivation/versioning processes in order to be made available for dissemination/reuse. There are similar steps in the digital content lifecycle, where new content from creators and/or distributors goes through archival processing and aggregation procedures in

order to be made available for access by users. Both lifecycles include steps of acquisition, processing, integration, and distribution. In some cases, creators and repositories may collaborate, seeing themselves as co-producers of a product designed to be available to both current audiences and future users, using the same media streams in both environments. It is the practices associated with these sorts of cutting-edge applications that we hope to document here.

Getting Started

General Resources and Requirements

The institutional resources required to support a project or program fall into three categories: human resources, equipment resources, and resources that support sustainability. Human resources include technical, production, and administrative staff. In many cases, human resources (most notably skills) are so dispersed within an organization that they are best harnessed when collected into multidisciplinary teams. This is not always the case. Policy-making, for example, is an issue that must be addressed at an institutional level, while gaining specific subject expertise is an individual role. However, project teams are especially suited to certain roles that neither individuals nor institutions can effectively fulfill. Collaboration with other teams, individuals, and organizations is one of the more notable of these.

For example, within the context of an institution-wide digital repository, collections of performing arts recordings may get lost in the shuffle if left under the aegis of the organization in charge of the entire repository. Similarly a single individual also may not be able to adequately provide for the maintenance of the collections. A team, however, is focused enough to devote detailed attention to such a limited domain, while at the same time being broad and varied enough to meet the correspondingly varied challenges of caring for the collections; able to form collaborative partnerships with the repository management, institutional legal counsel, contributing artists, and user groups, among others. Such team-based work may thus be seen as one of the requirements for effectively managing human resources.

Equipment requirements include all necessary physical facilities, devices, and other environmental factors, although this document will only focus on computing resources (e.g. networks, storage, content-management applications, etc.). Finally, funding and sustainability need to be taken into account. If the collection is built as part of an ongoing program—as opposed to a limited-term project—it needs to be able to carry on its activities for the foreseeable future. Hidden, diffuse, and ongoing costs—such as those incurred by loss of expertise through staff turnover, maintenance of physical facilities, and payment of staff salaries—need to be covered. The program need not necessarily cover its own costs, but these costs must be taken into account, and renewable sources of funding must be lined up. Even if the collection is organized as a project with a definitive end date, these issues cannot be entirely ignored. Moreover, some projects wish to evolve into programs. In these situations, project leaders should plan for (possibly) more complex costs, as well as ongoing financial support, as part of their transition. For more detailed information, see the ‘Resources’ and ‘Sustainability’ sections of the *NINCH Guide to Good Practice* [Introduction 6].

Resources Specifically for Asset Management

As mentioned earlier, digital assets, like other records of human activity, follow an information lifecycle. The resources and requirements of this lifecycle also need to be taken into account when planning repositories. For example, information about the performances and their recordings may not be considered an obvious budget item to those accounting for the resources of performance projects. Yet this information about the digital objects in the collection—this metadata—is a vital asset of any repository, which itself must be tracked as any other valuable resource. Metadata is used in every aspect of the information lifecycle. For example, metadata is used to identify creators of works and sources of objects, specify technical and operating environments, schedule preservation activities, facilitate search and retrieval, explicate rights and

restrictions to access, implement file playback, and track use. As such, systems for managing metadata are one of the special requirements of asset management.

Also, it is important to reiterate that collection management requires diverse people, equipment, and support at individual, team, and institutional levels. For example, preservation activities may require not only the proper supplies and technology, but also investments in security, redundant systems, etc. Digital Asset Management systems will likely require a staff knowledgeable in the administration of RDBMS, web and/or streaming media servers, and various XML metadata protocol and standard applications. Similarly, the implementation of online rights and access management systems—to track rights, licenses, and user authorization/privacy issues—will also require specialized applications and the personnel to manage them. Part of the purpose of this guide is to help readers identify possible points of collaboration and skills to look for in team members; to that end some discussion of many of these areas is found in these guidelines.

Current Practices in Institutional Policies

While this document provides guidance primarily for practices related to team-level issues, in many cases institutional policy decisions greatly affect team practice. Administrators, library systems personnel, and teams of faculty and staff involved in the archiving of digital collections of performance events should all make themselves aware of their parent institutions' policies related to electronic information. They may find their collections subject to many higher-level policies, grouped here into the following five areas: archive definitions, ingest and archiving, preservation, dissemination, access and use, and users. The Working Group believes that if people carefully consider the following questions in each of these categories, they can be prepared for potential problems in these areas that may arise mid-project.

Archive Definition

- What is the purpose of the institutional archive? How faithfully does this stated purpose support the mission of the institution on one hand and the needs of the individual collections on the other? Are there opportunities for teams responsible for specific collections to create 'sub-repositories' within the main archive that align more closely with their specific needs?
- Is there a core asset of the overall archive that must be maintained above all else? This could either be in terms of format (e.g. text, image, or audio resources) or in terms of function (e.g. a public archive with a main goal of accessibility—and thus uses copies—as opposed to a 'dark' archive, which would presumably focus more on preservation copies). This has implications for many subsequent decisions. As a simplistic example, if a team finds itself in an environment where the core institutional assets are written records, the parent repository may have chosen a metadata standard designed for text documents (like TEI or EAD). The team in charge of creating collections of digital performing arts recordings would then need to figure out how to get the repository to accommodate its audio and video files. On the other hand, if multimedia files are the parent repository's core assets, it likely will have already chosen a more suitable metadata standard, designed with multimedia objects in mind (like METS).
- Does the team represent a program or a project? Is there assembled funding, expertise and staffing resources to foster a sustainable effort or is there an end date to the team's efforts? If the team is temporary, what efforts have been made to ensure that the performance collections will be maintained? Will other groups take over management of the collections? How has any transition process from team project to repository program been planned?

- Who are the stakeholders (library/archives, performing arts administrators, media programs, artistic faculty and scholars, general users, etc.)? As corollaries, whose requirements will drive policy and who is ultimately responsible for the collection? The answers to these questions indicate optimum levels of involvement, consultation and consensus among stakeholders, including commitments of staff, funding, technology and other resources. Through their teams, scholars will want to be sure that their voices are heard when such decisions are made. Moreover, there should be a process to ensure that whomever is directly responsible for performance collections should also be accountable for, and can justify the decisions made, on all of these issues to the collection's stakeholders.
- Who keeps what? Within the parent institution, is there a central, institution-wide repository, or are holdings decentralized (e.g. are similar collections divided among a university's archives, library, TV station, various performing arts departments, etc.)? If holdings are decentralized, how much coordination exists and/or is desirable between the different repositories? Are there clear lines of responsibility/jurisdiction for what sorts of digital performance content end up where?

Ingest and Archiving

- What is the repository's collection/deposit policy? How open are the collections to ongoing/future accessions? What provisions are in place to ensure that open collections nevertheless maintain their individual identities despite the acquisition of new material?
- What formats are accepted? Are there "guarantees" that the archive will continue to make all acceptable formats accessible in perpetuity or does it only extend this assurance to the most popular formats? Will the repository accept unusual formats with the stipulation that it may not guarantee their future accessibility? If it does provisionally accept these "fringe" or "outlier" formats, does it make efforts to preserve them on an 'as-possible' basis or not at all?
- What metadata standards do the archive use and what are the preferred methods and levels of implementation? To what extent will the repository's metadata and cataloguing staff stick with elements provided by the creators/donors, and to what extent does it normalize metadata elements into standard formats (or require creators/donors to apply such standardization prior to ingest)? The answers to these questions may be dependent upon the answers to many of the questions in the "Archive Definitions" section.
- How does the archive handle rights management (both acquisitions and conferrals of rights)? Has it created its own template or draft licenses, does the parent institution have model licenses that the archive has adapted for its own use, or does the archive negotiate licenses on a case-by-case basis?

Preservation

- Given the repository's resources, what kind of risk management and disaster preparedness policies can it afford to have? Has it identified and arranged for acceptable levels and types of mirror sites, off-site storage (including out-of-state storage if necessary), insurance, ways to get in touch with emergency contacts, etc.? Can it afford to *not* have any of these sorts of policies?
- How does the institution sustain organizational memory within the archive? What sorts of procedures and policies are maintained to guarantee consistency through time, staff turnover and changes in technology? How are these procedures and policies communicated?
- How has the archive defined policy and procedures for data refreshing and format migration? Does it have business plans and scheduling for hardware and software replacement? If it doesn't have the resources to schedule regular technology migrations, what plans are in place to ensure proper migration when it takes place on an ad-hoc basis (e.g. in the event of computer failure or unworkably out-of-date equipment)?
- What has been planned for the worst cases? Has the institution created relationships, policies and procedures to hand the collections off to other organizations and/or entities in the event of various

doomsday scenarios? Even the most unlikely events need to be planned for, and if the parent institution and/or host repository is lacking in this area, teams need to implement plans for this or potentially risk losing everything.

Dissemination, Access, and Use

- How has the repository provided for both dissemination of and access to its collections? These are separate issues. A single batch of content can be presented (e.g. disseminated) in multiple forms, with different applications for different audiences (such as different age levels) or redistributed in different versions as content for other repositories and so on. For example, footage from a performance may be disseminated in its entirety as a live broadcast, inter-cut with backstage footage in a documentary on the performing organization, or excerpted and packaged along with other resources on the history of performance genres in a courseware module. Audiences might access the first two of these by watching television or packaged videos or DVDs, the latter through an authentication-required courseware system, or any of the above through online digital resource archives. ‘Access’ deals with the interface, through which the user can play back a resource, whereas ‘dissemination’ deals with the package in which the repository distributes it. In what forms have the creators disseminated the content, are there provisions for repurposing the material, and if so, who is allowed to do so (e.g. are users allowed to repurpose content or disseminate it in other forms)?
- How are permissions and rights management decisions implemented at the user level? What restrictions can the repository system impose on the collections regarding type of use? How nuanced and flexible are the systems that manage these restrictions? How are rights communicated to users, and how is repository (if not user) compliance to those rights tracked?
- Once it is established that users may access certain content, what sorts of mechanisms does the repository employ for content delivery? Is the team’s content such that certain capabilities must be (dis) allowed? For example, should users be able to download content or merely stream it? If they are allowed to download it, should they then be able to transfer that content into physical form (e.g. PDF files that allow the “print” function to be removed from texts, even downloadable texts)?
- Whether due to access restrictions or the desire to work with users’ technical limitations, how may the resources in the collections be viewed? Can they be streamed over 56K modems by professors from home or will they only be available through the intra-campus network? Are restrictions so stringent that there may only be onsite playback of the material?
- How can the collections be searched? Does the system provide for content-based searching (of low-level features within audio or video resources) in addition to traditional catalogue-based searching of the collections? Is cross- or multi-modal searching desirable or feasible? (E.g. can a digitized capture of a sonata be searched for by its melody in addition to by its composer or performer?)
- If the decision is made to keep users from accessing certain performances, is the distribution of summaries (manually and/or automatically extracted series of clips representative of the whole) of these recordings legally and/or technically feasible?
- Given the possibilities of different distributions of, multiple types of access to, and multimedia summaries associated with resources in the collections, how will the archive track multiple versions of a resource and relate them to each other? Does the repository have the technical infrastructure to support all of these different functionalities and types of use?

Users

- What kinds of users will access the finished product: researchers, students, the general public, professional artists, others? Of these, which group(s) is/are the primary focus of the archive? Of the

team managing the collection? How is any discrepancy between these primary user groups reconciled?

- What will users want to do with the recordings? How will they want to manipulate the recordings, and what sort of search/browse capabilities will they want, both across the collection as a whole and within individual works?
- How much second-guessing of users' needs and wants is permissible? For example, eliminating and/or lessening audience and other background noise through re-mixing may be requested by music theory researchers and not at all wanted by ethnomusicologists. Or both groups may prefer to have such 'primary source' evidence left unchanged so they can make their own judgments regarding the material.
- What are users' levels of technical expertise? How independent are they and what levels of support will they require? What computing platforms do they use, and what playback software are they likely to have? If the archive contains and supports more formats than users can (dis)play, what provisions does the archive make to ensure that users still have access to all of these formats?

Preservation

For many digital repositories, including collections of performance recordings, it is premature to outline "best practices" since both the technological and the legal/economic environments within which long term preservation will occur are changing, and many of the challenges remain unresolved. Moreover, it is probably even premature to identify a consensus on a number of issues. That said, guidance could be offered on a set of issues that would-be managers of digital content should be aware of as they ready their collections for long-term preservation. Teams embarking on the archiving of performance collections can use this guidance as a framework for considering how they may help their parent institution provide for the preservation of their collections.

It should be noted that many of the principles and considerations outlined below are also covered in other aspects of this guide. This apparent redundancy reflects the reality that "preservation" is not a purely technical issue but occurs within a set of institutional contexts. Also, it should be emphasized that strategies for long term archiving of digital content are a subject of research. The challenge for archivists and librarians is to save material in digital form while remaining ready to adapt to change. Creators must realize that the technical formats used during a performance may not be the ones best suited to preservation. Therefore, repository-based collaborators and artistic discipline-based collaborators need to work with each other to implement preservation processes that honor both the original artistic vision of a work and the commitment to the future existence of that vision.

Moving Toward Good Practice for Preservation in Archives of Digital Information

- (1) Preservation begins at creation. Digital works are frequently dynamic and/or have been created in a given technological context, which is subject to change. Thus, capturing key information to enable continued use of the item can only occur at creation, when metadata creation is also cheapest. The minimum set of key elements that should be captured at creation has not been determined for performances or other digital works. Metadata schema, including for preservation metadata, have been proposed (see for example, the work done by the National Library of New Zealand). Thus, determining appropriate preservation metadata is key. While in the long-term, this is an area where artists, researchers, and repositories need to work on defining appropriate standards (as in [6], on metadata schema, below), in the short-term all participants should work to

ensure that metadata capture is systematized, consistent, and built into the earliest stages of production.

- (2) Decide on architecture. The repository has to have some sort of organization that consists of components. The architecture specifies the components, their boundaries, contents, etc. A number of architectures have been proposed (e.g., OAIS), but this remains an area of research. (See the section “Metadata Standards, Schemas, Frameworks, and Formats” for more information.) To get started, at a minimum, architecture should be specified. Among the considerations are:

- How will items be registered?
- How will the obsolescence of the repository architecture itself be managed (e.g., if the repository is built using a commercial database platform, what happens if that platform is changed)?
- Are items managed individually, within collections, or both?
- Where are the metadata stored – bundled with the work, separately, or segmented?

At a team level, groups working on building collections of digital performance recordings may not have complete control over the choice of repository architecture, because they may need to work within the framework chosen by their parent institutions. Nevertheless, the answers to these questions impact the long-term viability of the work and, therefore, teams should work with policy-makers to ensure that their needs in this area are being met.

- (3) Digital performance works can be dynamic. Among the considerations are:

- Who decides the conceptual boundaries of the work and determines what is preserved?
- How is that decision recorded and made visible for future users/managers of the archive?
- For works that are independently bounded, but conceptually linked (e.g., successive performances of the same program with different performers), how are these relationships preserved?
- Again, this information may be recorded as part of the identifier or the metadata, and there may be layers of metadata (collection level, work level, etc.). How is this decision made?
- How is the decision communicated to future users?

This is an area where collaborative teams can have particular influence. Such teams allow repository managers and technical specialists to work directly with the artists to determine what, precisely, comprises the work and what is necessary to maintain its integrity.

- (4) Individual works exist in a hardware and software environment. The relationship between the work and that environment, along with the use of the work, determines the appropriate preservation strategy (i.e., migration, encapsulation, emulation). A determination must be made about the correct strategy for the works in the collection. This is very likely to affect the metadata. Researchers and artists on teams can help repository team members determine this strategy by explicating the use(s) to which the works will be put. This is especially true in cases where the works will be part of a system intended only (or primarily) for the use of the creators and/or their academic department, with limited or no outside/general use. In these cases the creators may be able to predict the type and amount of use to anticipate for the works in the repository.

(5) Formats affect preservation decisions. Images, moving images, sound, text, etc. are all likely to require different treatment. For very large files, for example, will the master consist of an uncompressed file with compressed distribution versions for use? Is the distribution version the concern of the archive or only the master? What appears to be the “industry standard” for the format in question? Given the current state of the technology, there is a trade-off between the most innovative software tools, which are the most subject to change and potentially most interesting to the performing arts community, and the formats that are well understood and most suited to preservation. Indeed, some streaming media are ephemeral and projects that involved these media might have to be captured in an alternative medium, just as ballet is presently captured through still photography and film. Like metadata capture, this is another issue that teams must take into account from the very earliest stages of the production process. While there is much appeal to the idea that technology should adapt itself to users (in this case performing artists), rather than vice versa, it is also true that digital repositories are constrained by the limits of today’s technology. By taking format factors into account from the beginning, artists can help repositories ensure the long-term availability of their works.

(6) Choose a preservation metadata schema. Defining preservation metadata is an active area of research, but to get started (as is the case in [2], regarding repository architecture, above), a decision about preservation metadata must be made—in full recognition that any decisions will have to evolve as the state of the art advances and consensus emerges. Therefore, any schema developed must be characterized by flexibility, extensibility and—most importantly—a commitment to change. Several key organizational issues are also associated with preservation metadata:

- Who is responsible for assigning the metadata?
- Who is responsible for archiving the metadata?
- What identifier scheme is selected? (Identifiers should be unique and persistent.)
- Who and what entity is responsible for naming/assigning identifiers to works?
- What happens if the identifier scheme is rendered obsolete?

Again, the earlier these decisions are made and the more that all participants are aware of and involved in the decision making process, the more likely it is that the decisions will be implemented satisfactorily and applied consistently to the works.

Intellectual Property

This section deals with the intellectual property issues that are likely to arise when dealing with digital collections of live performance events. It starts off with a hypothetical scenario to provide an example highlighting the interactions of the many intellectual property concerns of digitally captured performances, and then reviews certain facets of intellectual property in relation to performance collections. While all collections (digital or otherwise) require attention to intellectual property issues, research and educational performance collections have four characteristics that make their management especially complex:

- The objects in the collections have many creators, and indeed many types of creators, each with differing types of rights.
- Some of these different types of rights holders are among the users of the collections, and therefore may be authorized to have access to the materials beyond that deemed ‘fair’ under the fair use exceptions of various copyright acts.
- Because these collections are the results of research and educational activities, they may require different (higher) levels of accountability due to either Institutional Review Board policies or student record privacy laws.
- Performance collections, which document cultural heritage and patrimony, may also require awareness of and sensitivity to moral and ethical issues of cultural appropriation and representation, especially when the cultural background of those involved with the capture, research, and/or archiving of these collections differs from that of the performers and/or artistic creators of the works so documented.

All team members—whether from artistic, repository, or technical backgrounds—need to be aware that repositories are responsible for the application of these issues in their collections, a task that does not necessarily get easier in the digital realm. When implementing licensing and rights agreements for performance works, repositories that hold collections of recordings of events must be sensitive to the needs of the works’ creators, performers, and users. Moreover, the repository needs to take into account cases where these categories overlap. For example, if an academic institution captures and archives its own student and faculty performances, it may have to interact with its students and faculty as both creators from whom it licenses the rights to digital recordings and as users to whom it provides access to those recordings. In the following scenario these interrelationships are taken to an extreme in order to present a complex situation that highlights the possibilities (and potential pitfalls) of managing intellectual property in collections of digitized performance.

- A professor in a theatre department is using the digital video of a play he or she directed last year as a resource for a senior honors theatre class this term.
- A few of the students in this class were also actors in the play.
- The script of the play was written (to fulfill a degree requirement) by the professor’s teaching assistant, a graduate student in the department’s Dramatic Writing Program. (Presumably, because it was for credit, it is also a part of his or her student record.)
- The costume, set, lighting, and prop designs were created by a mix of student volunteers and contracted professionals; again a few of the student designers are in the class.
- The soundtrack for the play was a selection of popular songs; when the play was produced the department obtained the rights to use the songs in the play during its original performance run, but not for subsequent broadcast and/or distribution.

- One of the professor's colleagues, a professor in the film department, originally was in charge of the recording of the performance. This professor now also wishes to use the digital video as a class resource, in a practicum on digital video editing and production.
- Because the students in this class will be working with the raw audio and video streams from the capture, each in effect will be creating a new edition of the recording of the performance.

These kinds of complexities make it a requirement that custodians of academic and/or research performance collections keep close watch over the management of rights, licenses, and privacy controls. Because these collections originate in academia, which values the open exchange of ideas, this implies a commitment on the part of their creators and archivists to maintain adequate provisions for fair use—even while all of the collections' other intellectual property needs are being met. The following sections provide further overviews of these topics.

Rights and Licensing

This covers both the familiar notion of copyright and more esoteric rights that creators of an intellectual work may reserve. However, not all types of creators will be able to claim each of these kinds of rights. Indeed, several groups will likely be involved in the creation of the digital objects. Creators of the intellectual work being performed include writers, composers, choreographers, and designers. All may hold copyright over their individual contributions to the final performance. Similarly, if those responsible for the recording of the performance made a significant contribution to the finished recording beyond simple capture (e.g. significant editing, the creation of a narrative, inclusion of other works, etc.), they may also be able to claim a separate copyright for that recording. Depending on the needs of the project, certain creators may ask others (with a smaller role in the creation of the work) to waive their claims to a work by signing a release. When placing performance resources in a digital repository, creators and aggregators alike need to think about how they will document and track the releases they've obtained and the rights they've cleared for the work in question.

This can be complicated, especially when performances incorporate other works. For example, a dance or theatrical production may use copyrighted works as background music. If the producers of the recording neglected to obtain additional distribution rights along with the right to play the music in the original performance, the repository may not be able to webcast or otherwise replay that recording. There are myriad types of rights relating to any work, including distribution rights (such as for broadcast or web streaming), the right to alter the work or create derivative works from the original, and the right to access the material. Alteration rights are discussed more below, in relation to licensing.

Access/render and distribution rights differ from copyright in that they control the terms under which users may view the material, irrespective of any reproductions that may be made. For example, in the digital realm the accessibility of a file encoded in a particular format may be dependent upon the availability of a third-party proprietary application for rendering that file format. In such a case, if the creators of the file deemed the format unique and valuable enough that migration to other formats was not desired, they could ensure the accessibility of their file by entering into an agreement with the third party. This agreement would guarantee that the creators continued to have the right to access the proprietary viewing application even in the event that the third party was unable and/or unwilling to provide support for that application. Such an agreement might decree that the third party put a copy of the code of the application software

into escrow, so that if such an event occurred, the application would be released to the creators of the file. However, many creators will not be able to apply for and/or receive these types of licenses. In this case, they must be aware that esoteric, cutting edge formats may not be re-playable if repositories are unable to themselves get licenses for the software needed to render those formats.

Distribution rights are somewhat similar to access rights, in that they do not specify whether a work may be reproduced in addition to being made available. However, instead of ensuring creators and/or repositories the right to be able to hear or view the work despite changes to its underlying encoding, distribution rights grant repositories the right to make works available to their end-users, whether by broadcast or publication, and whether via traditional media or over the Internet. This area is of special concern to teams that wish to make performance recordings available via the web, whether as part of online research and courseware environments or in a publicly accessible digital collection. Because webcasting is such a recent phenomenon, the legal and contractual requirements for clearing such distribution rights have not been fully defined in either government regulation or common practice.

Generally, the licensing process determines what the rights-holders of a work allow users to do with that work. In the case where the work is collected in a repository that then offers that work to its own users, licenses can be both between the creators and the repository and between the repository and its end-users. For example, the license between the rights-holder and the repository specifies whether the repository is allowed to charge its users fees and/or it is required to pay the rights-holder royalties for using the work. The same repository may then have a license agreement with its end-users that specifies that the users agree to pay those charges. In some cases the repository is the rights-holder for the work in question, which removes one layer of complexity in the administration of the rights. Nevertheless, in all cases, the rights-holders have the final say over who gets to use which works in what manner (except when the use is deemed to fall under 'fair use,' see below).

Examples of rights that may be itemized in a license also include the right to reproduce or alter a work. For example, rights-holders may grant a repository the right to make limited archival, preservation, or backup copies of a work in order to help guarantee the repository's right to access the work and help forestall deterioration due to use or outright loss. At the same time, the rights-holders might forbid the repository to let its end-users make digital or physical reproductions of that work by downloading it or printing/re-recording it onto physical media. Alternatively, the rights-holder may indicate in the license that the repository only extend such rights under certain conditions, for example serious researchers the right to high-quality versions of a work as long as the researchers contact the rights-holder for final approval, or granting all users the right to download lower-quality thumbnail versions of the work. Of course, for the rights-holder to let the repository make lower-quality versions of a work implies that it has given the repository some right to change the original digital object. Questions relating to what sort of changes can be made to a work fall under the category of alterations, in which the issue of thumbnails is but one. For example, are other types of summaries and abstractions of a work, such as clips of sound recordings or screen-shots of videos, allowed? What about alterations intended to improve the quality of a recording, such as preservation re-mastering, noise reduction or color correction? Additionally, in the case of alterations intended to help preserve the original, does the repository need the permission of any or all of the rights-holders to make these sorts of

changes, or can it go ahead and implement these processes as circumstances (and perhaps its users) demand?

This last question highlights one of the differences between physical and digital objects. Physical documents, such as books, generally fall under the doctrine of first sale. First sale expresses the concept that people who buy a legal copy of a work have the right to do whatever they want to that copy, provided it doesn't violate copyright. For example, the owner of a book can rebind it in order to make it last longer, make notes on its pages, or sell it to another person, thereby allowing that second person the right to do these things. A license to a work, however, is not a purchase of that work. Many licenses in the digital realm specifically prohibit actions analogous to the book scenario being enacted upon the digital work covered by the license. While some licenses allow alterations to be made to a work for the sake of preservation, it is unusual to find licenses that allow users to annotate or otherwise embed their own commentary within digital works. Moreover, some licenses specify that the licensee may not transfer the rights granted to it to any other entities. Among the many issues creators and repositories must confront is whether, and how, to maintain first sale provisions in their licenses.

Privacy

Privacy concerns primarily regard the rights of performers of works documented in the collection. While performers are not usually able to obtain copyright for their performance (with certain notable exceptions), they may be protected under privacy laws and regulations. The policies of individual educational institutions' Institutional Review Boards (IRBs) may be a matter of concern in some collections of performances used in ethnographic research or performances that include minors; also of concern is the Family Educational Rights and Privacy Act of 1974 (FERPA) [Intellectual Property 3], which regulates the use of and access to student records.

IRB policies, which govern the rights of human subjects when they are used for research, tend to be a more obvious concern of medical and psychological research than of artistic or cultural research. However groups and individuals capturing performances are not exempt from IRB review; this sort of inquiry will generally fall under the same sort of scrutiny as other types of ethnographic or anthropological research. Moreover ethical and scholarly principles need to be adhered to, in addition to legal and regulatory obligations. For example, many issues relating to appropriation of a group's cultural heritage, the rights to national patrimony, the ownership of folklore, etc., have not been fully resolved.

FERPA gives students the rights to view and amend their student records, as well as some control over how and to whom those records are disclosed. FERPA regulations place strict limits on the dissemination of personally identifiable information, i.e. information with which people might identify or locate students without their consent. Usually this means name, address, contact, and other demographic information. However, because video, especially, captures information about student performers that can help identify them (e.g. their pictures), repositories need to take care to accommodate FERPA regulations if they cannot be sure that the students in these recordings have waived control over their access.

Also, there may be uncertainties about some items in these collections in regards to publicity rights. Publicity rights are related to privacy rights in that they allow people (and more relevantly to this discussion, performers) the right to determine how their image and/or

likenesses are used. While the right of publicity is most applicable to celebrities, it may come up when famous performers work as guests or artists-in-residence [Intellectual Property 5]. Nevertheless, institutions that capture work may wish to include both privacy and publicity rights in their release forms. In cases where these release forms indicate that performers/participants may specifically exclude the assignation of certain rights, collecting institutions may need to check for publicity rights in addition to more familiar kinds of rights. Jill Gemmill from the University of Alabama at Birmingham has compiled several examples of release forms and put them up on the ViDe website, at <http://www.vide.net/resources/forms/>.

In each of these cases, responsibility for protecting the rights of performers falls to both the creators of the recording and to the digital repository. Creators are likely to be responsible for negotiating IRB approval and obtaining proper releases and consent forms (either themselves or in collaboration with the artistic production team), while repositories may be responsible for double-checking with content creators to ensure appropriate care was taken at the outset, and/or whether access to certain recordings will need to be restricted in order to maintain the privacy or respect the cultural patrimony of the performers. Both capture teams and archiving teams should consider how they will approach these areas, including what sorts of collaborative partnerships they need to make in order to implement their chosen approaches.

Fair Use

The fair use exceptions and exemptions are portions of the copyright code that decree that certain uses of works either do not violate the copyright of those works, or do so in a ‘fair’ manner. In effect, fair use gives users the right to quote, adapt, or otherwise use portions of a work. Many academic institutions have drafted institution-wide statements regarding fair use, generally through collaborations of the library system and the institution’s legal counsel. [For a few examples of guidelines from colleges and universities for determining whether a use is fair, see Intellectual Property 1, 2, 6, and 7.] When placing performance resources into an institutional repository, rights holders should give thought to whether any specific resource merits giving certain audiences/users either more or less extensive fair use rights than are normally provided by the general fair-use policy. With these decisions in place, statements explicating any notable restrictions or allowed uses should be incorporated into the license for that resource, and tracked in its metadata.

Metadata Standards, Schemas, Frameworks, and Formats

Metadata is information about resources. It is often used to create representations of, and surrogates for, resources such as finding aids, catalogue records, header and other ‘about’ files, citations, stylesheets, summaries, and indexes. These representations and surrogates in turn allow resources to be identified, categorized, searched, retrieved, displayed, and generally managed. While all metadata describes some aspect of a resource, either its content, its maintenance, or its use, metadata elements can also be sorted into three broad categories: descriptive, administrative, and structural metadata.

- Descriptive Metadata is data about resources analogous to the bibliographic information about books found in library catalogues (indeed, bibliographic information is descriptive metadata about print resources). Examples of descriptive metadata include title, creator, publisher/distributor, subject, date of creation or release, and physical characteristics.
- Structural Metadata is data that represents how the work is put together, or how parts of the work relate either to other parts or to a larger whole. Examples of structural metadata include table of contents with page numbers in print resources, clip structure linked to timecode stamps in video, the of a web document, and sequence structure of multimedia resources in a courseware object.
- Administrative Metadata is data that helps those currently responsible for the resource to manage its use. Examples of administrative metadata include information about the original technical environment and requirements of a file, provenance or conservation history, acquisition number, retention schedule, access restrictions, royalties due. As of yet, the standardization of administrative metadata has not progressed as far as the standardization of structural, and especially, descriptive metadata. In terms of technical metadata for multimedia, the most notable standard is the National Information Standards Organization (NISO) IMG standard for technical metadata for images [Metadata 18]. Research into standards for other types of administrative metadata is even more preliminary. For example, while a few groups have proposed XML-based languages for digital rights management, this is an area still in need of much study before the schemes become standards ready for implementation.

Metadata standards often mix elements from each of these categories to form a more workable tool. For example, MARC, while a descriptive metadata standard designed to encode card-catalogue information, includes tags to encode both administrative and structural elements. Also, standards grow out of different communities with varying design emphases. For example, some communities consider standards that encode detailed technical information about a digital resource (e.g. the type of computing environment needed to run/display the file) to be administrative metadata, while others consider technical metadata its own category.

The following sections include both emerging standards specifically designed for the web and traditional bibliographic standards that have migrated to the online environment, and each has its own idiosyncrasies. Altogether, a way to categorize metadata standards other than by type of data encoded can sometimes be more helpful. In this document, standards are organized according to the granularity of the information they are meant to record and include metadata standards at many different levels of granularity, presented in the following order:

- Reference frameworks that describe the overall organization of content delivery systems and repositories,

- Structural encoding schemes for multimedia and/or complex objects
- Standards for the encoding and transmission of once-paper catalogue records
- Core element vocabularies to provide similar (albeit simplified) discovery metadata for digital assets,
- Subject heading lists and thesauri to standardize the keywords used within individual descriptive elements.

Some Commonly Used Metadata Schemes

Helpful to a discussion of metadata standards is a brief digression into Extensible Markup Language (XML) [Metadata 26]. XML is not a metadata standard per se, but the understanding of many of these digital collection standards is greatly enhanced by a passing familiarity with XML. Like HTML, XML is an encoding scheme derived from the Standard Generalized Markup Language (SGML) [Metadata 25, 27]. However, while HTML is an SGML language used to encode web pages, XML is actually a family of related meta-languages used to encode other markup languages. These languages in turn can be used to encode almost any sort of information. XML languages described in this document include METS, MPEG-7, MPEG-21, IMS, SCORM, and MODS. Also, the Dublin Core Metadata Initiative is working on a direct XML implementation of Dublin Core. Other XML languages used in library/archival settings include RDF (Resource Description Framework, which has been one of the preferred means of implementing Dublin Core in XML), and SMIL (Synchronized Multimedia Integration Language, used to create multimedia presentations and applications) [Metadata 28, 29]. Additionally, certain SGML languages that had been created before the emergence of XML, such as EAD (Encoded Archival Description) and TEI (Text Encoding Initiative), now have XML versions [Metadata 6, 21]. Moreover, even metadata standards that were created and are now in use independently of the SGML family of languages (e.g. MARC) have gained XML implementations (specifically MARCXML). Updated information on most of these languages, as well as many other XML developments and initiatives, can be found on Robin Cover's directory of SGML and XML languages, the Cover Pages [Metadata 4]

It is important to note that organizations may use in-house metadata schemes customized to their own needs instead of, or in addition to, any of the standards below. While the use of standards helps assure interoperability of metadata, not all projects require this attribute. Moreover, careful design work on an in-house scheme can provide a measure of interoperability while meeting the need for precise customization. For example, when creating a metadata schema to use for resources encoded in its Making of America 2 format (the forerunner of the METS standard described below), the University of California at Berkeley created the Generic Descriptive Metadata scheme (GDM) [Metadata 22]. GDM specifies descriptive metadata at a level of detail in between those provided by the Dublin Core and MARC standards (both described below), letting GDM users describe their resources more fully than would be possible with Dublin Core, without all of the extra effort required by MARC. While Berkeley and other implementers of METS are now considering replacing the use of GDM with the recent MODS standard (also described below), other repositories may find themselves in similar situations, especially if they must uniformly describe content received from many different creators (and other sources), each of which may employ its own metadata scheme.

Project teams should examine the relevant standards, as well as any in-house specifications, and decide which metadata schemes best fits the needs of the project in question. In situations where the creators of the performance recordings are part of the same institution as the collection's repository, those overseeing the production of the recording should be made aware of the metadata scheme used by the repository. This way, they will know what metadata they need to collect at the time of capture, and can make this process a step in their planning. This is another area where the project team may benefit from the inclusion of creators. Not only will the metadata requirements be more effectively communicated to those doing capture if they are part of the team from the beginning, but their insights into the artistic production process will help more repository-oriented collaborators know what information about digital recordings the artists find it important to track and keep.

Repository Frameworks and Structural Encoding Standards

METS: The Metadata Encoding and Transmission Standard (METS) is a standard maintained by the Library of Congress for encoding descriptive, administrative, structural, and behavioral metadata regarding objects within a digital library [Metadata 15]. Compared to many of the other standards described here, METS is especially concerned with the structural representation of a resource. METS' emphasis on structural metadata and its capacity for very complex logical and physical structures allow it to represent many disparate genres of resources. So, for example, an entire run of a broadcast series can be described with a logical hierarchy (series, episode, segment, scene, clip, shot, etc.), while a theatre scrapbook could be described through a physical hierarchy (covers enclosing pages, many of which contain text, while some also contain attached newspaper clippings, while others may contain glued-in envelopes with costume samples or music CDs inside, and so on). Also of interest is METS' support for behavioral metadata, the encoding of which allows reactions and behaviors to be associated with the resource in response to specific user actions. METS is designed to work with the Open Archival Information System Reference Model (OAIS). OAIS is an ISO international standard for archive systems for digital information [Metadata 3, 19]. As a framework for maintaining the long-term preservation and accessibility of such information, OAIS does not specify specific implementations. Instead, it describes repositories' responsibilities and the conceptual systems architecture needed to coordinate activities and the transfer of digital objects between producers, repositories, and users. OAIS defines its digital objects, Information Packages, as:

conceptual container[s] of two types of information called Content Information and Preservation Description Information (PDI). The Content Information and PDI are viewed as being encapsulated and identifiable by the Packaging Information. The resulting package is viewed as being discoverable by virtue of the Descriptive Information. [Metadata 3]

In this model, Descriptive Information is descriptive metadata, while Content, Preservation Description, and Packaging Information are mixtures of administrative and structural metadata.

MPEG-7: As part of its work to create ISO standards for multimedia resources, the Motion Picture Experts Group (MPEG) has also produced a standard for multimedia metadata, the Multimedia Content Description Interface [Metadata 11]. MPEG-7 is a standard for describing multimedia content in such a way as to make it searchable at many levels. MPEG-7 'Descriptors' can be used to represent content features from the very fine-grained (like color histogram or tempo information) to the very abstract (like the genre or subject of a work). This is meant to facilitate retrieval of the described content on any level that the search application (not specified as part of the standard) allows. In addition to content metadata, MPEG-7 descriptors also cover

other descriptive (title, director, encoding), administrative (copyright, usage history), and structural (location and type of scene cuts, spatial sound representation) metadata. The Motion Pictures Experts Group has also started working on MPEG-21, the Multimedia Framework [Metadata 12]. In some ways, MPEG-21 is even broader in scope than OAIS, focusing on the architecture at the ends of the production chain (producers and consumers) rather than in the middle (the repository). As the MPEG-21 documentation notes:

The scope of MPEG-21 could be described as the integration of the critical technologies enabling transparent and augmented use of multimedia resources across a wide range of networks and devices to support functions such as: content creation, content production, content distribution, content consumption and usage, content packaging, intellectual property management and protection, content identification and description, financial management, user privacy, terminals and network resource abstraction, content representation and event reporting. [Metadata 12]

This vision de-emphasizes the role of repositories, seeing them as just one link in the multimedia delivery chain. This may create concerns about these standards for those who are in repository settings. Unlike many of the other standards here, the MPEG standards do not come out of the library/archival world. Instead, they come from MPEG, which is largely a group of industry experts from the domains of broadcasting, film and music studios, computer and entertainment equipment manufacturers, etc. Therefore, they are neither as aligned with nor as responsive to the needs of repositories in academic and/or cultural heritage institutions. Moreover, both of these standards are very new. Many parts of the MPEG-21 standard are not scheduled for completion until 2003 or 2004. MPEG-7, while complete, is represented in fairly few implementations. The standards are mentioned here largely because they may become the basis for new commercial file and media formats around which collections may be based--especially if they follow the trajectory of earlier MPEG standards, such as MPEG-1 (which brought mp3s and video CDs) and MPEG-2 (which brought DVDs).

IMS: The Instructional Management Systems (IMS) Global Learning Consortium has produced a number of XML-based specifications that address various aspects of courseware management [Metadata 10]. These systems provide another way of thinking about and using multimedia resources for research and education. Of particular interest are its existing specifications for content packaging and metadata, and its specification under development for digital repositories. Like METS, the IMS Content Packaging specification is meant to encode the structure, location, and description of multimedia resources, albeit specifically resources integrated into a courseware system. The IMS specifications work together. The Content Packaging specification uses the IMS metadata specification in its description of resources; when the Digital Repository specification is complete, it will (similar to OAIS) provide a system for storing and managing collections of objects encoded with the Content Packaging specification. Many organizations have adopted and extended the IMS specifications for use in various commercial, academic, and governmental training and education situations. For example, the Advanced Distributed Learning Initiative created the Shareable Content Object Reference Model (SCORM), as a metadata system to encode web-based learning objects [Metadata 20]. (Originally SCORM stood for "Shareable Courseware Object Reference Model".) Originally founded by the Department of Defense, Advanced Distributed Learning is a consortium of public and private organizations working to unify disparate learning and course content management systems, standards, and specifications into a single content model. As such, SCORM incorporates IMS specifications and various other standards to present a flexible, high-level framework useable in a wide variety of distributed learning environments.

Standards For Encoding Metadata for Resource Discovery

Dublin Core: Dublin Core is a relatively widely adopted basic set of metadata fields [Metadata 5] comprising an extensible set of 15 elements from the Dublin Core Metadata Initiative. Dublin Core's elements (Title, Creator, Subject, Description, Publisher, Contributor, Date, Type, Format, Identifier, Source, Language, Relation, Coverage, and Rights) consist of a minimized set of descriptors in order to facilitate use and interoperability across a wide range of disciplines and industries. Because it is meant to provide the minimum amount of information necessary to locate and retrieve an object, the Dublin Core element set mixes the three types of metadata in order to produce a record that contains only the metadata absolutely essential (hence 'core') to the object. ANSI has recently (2001) approved Dublin Core as a standard and it is becoming a common means of translating between more detailed and nuanced metadata schemes. Other sets of core elements, like the Visual Resources Association's (VRA) Core Categories, function in much the same manner (albeit in ways customized for specific audiences) [Metadata 23]. The VRA Core is a set of 17 categories used to describe visual works and the digital images that document those works. While the VRA Core categories (Record Type, Type, Title, Measurements, Material, Technique, Creator, Date, Location, ID Number, Style/Period, Culture, Subject, Relation, Description, Source, and Rights) are similar to the Dublin Core elements, VRA Core allows a much more detailed representation of the physical form of the object being described than Dublin Core does.

MARC: Machine-Readable Cataloguing (MARC) is used to standardize the exchange of 'card catalogue' information in online public access catalogues, and provides a framework for describing a mixture of metadata [Metadata 14]. Like Dublin Core, its fields include descriptive (title-245, topical subject headings-650), administrative (cataloguing source-040, restrictions on access-506), and structural (computer file characteristics-256, organization and arrangement of materials-351) qualities. MARC enjoys widespread use in libraries, and is also fairly common in archives. Indeed, while MARC was originally meant only to catalogue individual items, at the behest of the archival community the format has been expanded to allow single MARC records to describe resources at any level of aggregation, up to and including collection-level description. However each MARC record can only describe a resource at one level; in order to provide description of a collection both at aggregate and item level, multiple MARC records would need to be created. These can be linked, but even when describing individual items the MARC is an arcane format that requires more specialized and extensive training to use than Dublin Core. The MARC standard is maintained by the Library of Congress, which also maintains MODS, the Metadata Object Description Schema [Metadata 16]. MODS is an XML language that can be used to translate MARC records into XML and/or to create native XML documents that encode MARC-like records. The tags used in MODS are a subset of MARC tags, and are language-based instead of numerical. As such, MODS documents are meant to be simpler to encode than full MARC records, but to provide richer descriptions than those provided by Dublin Core.

Content Standards

LCSH: The Library of Congress Subject Headings (LCSH) deal primarily with descriptive metadata [Metadata 13]. Like all controlled vocabularies, LCSH consists of an authorized terminology meant to avoid the problems of synonyms and multiple meanings, in this case in the terms used as subject headings for material in the Library of Congress. LCSH terms are often used in the subject fields of MARC and other catalogue records. Other subject heading

vocabularies, such as the Getty Museum’s Art & Architecture Thesaurus (AAT), provide for different ways to describe a variety of intellectual works [Metadata 7]. The AAT gives a considerably fuller means to describe objects and the built environment, allowing categorization not only of the subject of a work (what it is about), but also its form, genre, and physical characteristics (what it is an example of). Moreover, the AAT is faceted, with parallel hierarchies not only for physical characteristics and subject content but also for functions, time periods and styles, and the entities involved with an object or resource. This greater functionality means that AAT can be used not only in metadata records’ subject fields, but also in the fields describing these other characteristics.

Too Many Standards?

As may be gathered from even the very selected list above, there are myriad standards, schemas, formats, and guidelines for encoding metadata for digital objects. When choosing which standard or standards their institution should implement, project teams must be mindful of what best suits their mission and collections. Each scheme carries with it a certain level of complexity and power it can support in terms of the detail of description and the intricacy of associations it can encode. Moreover, this complexity and power comes with equivalent trade-offs in terms of the resources it costs to implement and maintain each scheme. Metadata implementations require the coordinated work of many stakeholders—including content creator commitment to metadata collection at time of capture, repository staff time to properly track and maintain metadata associated with digital objects, and IT support to integrate and maintain the server, network, and software environments—in order to be successful. If the chosen schemes are such that they require changes in work practice and/or extra effort in order to harness their complexity, all players must be made both aware of and prepared for the burdens that accompany the benefits of the new system.

Choice of standards is somewhat guided by the fact that some standards are meant to complement others. Indeed, with such wide variety in the level of metadata granularity that the standards specify, standards creators often look for ways in which their standards can interact with pre-existing schemes. METS, for example, was designed as a wrapper for digital objects consisting of resource files and their associated metadata; it is expected that one or more of the administrative or descriptive standards (NISO IMG, Dublin Core, MODS, etc.) will be used to structure the descriptive and administrative metadata associated with METS resources.

At a more granular level, content standards provide the vocabulary for descriptive elements within resource discovery standards. As mentioned above, LCSH is used in Dublin Core and MARC, and AAT can provide the vocabulary for many VRA Core elements. Similarly, standards complement each other at more overarching levels as well. For example, OAIS and MPEG-21 are meant to serve as overall frameworks for repository structure and practice and, as such, work with encoding and transmission schemes. For example, within an OAIS environment, METS documents could serve as the Information Packages encoding (and/or pointing to) the repositories’ digital objects and those objects’ associated metadata. MPEG-21, meanwhile, is meant to serve as a framework to guide the creation and distribution of multimedia objects encoded with its more-familiar cousins in the MPEG family. Teams intending to adopt additional metadata standards should make interoperability with other standards currently implemented in their organizations a criterion for examination and eventual use.

In addition to direct interoperability through complementary use as described above, another way to help guide these decisions is by examining the efforts taken to translate a particular metadata standard to other such schemes. Metadata maps, crosswalks, and harmonization projects also relate schemes to each other and allow for (and maintain) interoperability between groups who follow the standards involved. For example, the mandates of some countries' national standards bodies have the weight of law; if these bodies favor a particular standard, users within those countries must adhere to that standard. Crosswalks and other translation devices must then be created when users in other areas do not follow that standard, either because their countries favor different standards or because they do not mandate the use of any particular standard. Moreover, these efforts indicate that each of the schemes linked is used and supported, at least enough to warrant translation from one to the other and back. Schemes with many such linking efforts, such as Dublin Core and MARC, have proven that they have attracted widespread interest. For example, Jane Hunter, Jose M. Martinez, and Erik Oltmans have written about the links being made between MPEG-7 and Dublin Core in their paper "MPEG-7 Harmonization with Dublin Core: Current Status and Concerns" [Metadata 9, also see Metadata 8]. For a broader, more generally applicable set of crosswalks see the Getty Museum's "Metadata Standards Crosswalks" [Metadata 2]. Another list of crosswalks, as well as a primer and other resources on metadata, is available from the National Science Digital Library [Metadata 17].

Search and Retrieval

It cannot be emphasized enough to communicate with the university information technology groups throughout this process to learn about formats, standards and tools that are required or available. As with the formats and standards, the tools to search and retrieve provide an equally wide array of options. The simplest example is a web page that lists the performances similar to a catalog, with each performance listing a description and a URL that can be clicked to view the performance. The most complex is processing the performance through “scene-detection” software that identifies start and stop frames of various scenes and prompts the user to provide information (metadata) about the scene. The processed performance is loaded into a search engine that can retrieve and display the entire performance or specific clips that match search criteria.

The more sophisticated the indexing process, the more sophisticated and granular the searching and retrieval. For example, a user who wishes to find information regarding real-time motion-capture in live dance accompanied by computer-enhanced video recalls seeing video of dance/motion capture hybrid performances at a conference a few years ago, and therefore decides to search recently published dissertations. For dissertations that identify the author, advisor, university, and discipline in the index field, the user can search at a macro level and will eventually have to look at several different dissertations to find the correct one and preview that dissertation to find the correct scene. For dissertations that additionally index key subject terms and various scenes in terms of what appears in the scene as well as what the scene represents, a user can precisely search for “real-time dance motion capture live performance” and be presented with the exact scene from the dissertation.

It is, of course, possible for a performance to be indexed many different times through its life. It may initially be published with very simple indexing in a “catalog” website. Then as its importance or popularity grows it may be re-indexed with scene detection and published in a searchable digital library.

Digital Collection Software/Systems

As in the previous section, this section requires significant communication with the university IT groups responsible for hardware and software systems that store, search and retrieve digital assets. Again, there is a wide-array of options including accepting what the university has to offer, purchasing commercial applications, using open-source applications or outsourcing these issues to an Application Service Provider (ASP). The simple approach is to place the file in a directory and have a URL retrieve the file. As more files are created the greater the management becomes. Providing different formats to Microsoft versus Apple users or high-bandwidth versus low-bandwidth users starts to complicate the process of acquiring, converting, managing, searching and retrieving these assets.

To address this growing problem, several open source groups and commercial companies have offered products called Digital Asset Management (DAM) systems. These systems provide an “all-in-one” approach to managing digital assets. They acquire and convert the video input files, commonly in a batch mode, and can produce any or all of several different output files. They store all the files in a repository with associated metadata. Often they provide searching and retrieving capabilities. Some even provide “out-of-the-box” e-commerce solutions. There are several different, yet overlapping, technologies on the market that can provide sufficient functionality depending on your requirements. The following table summarizes some of them.

Technology	Acronym	Purpose	Providers
Digital Asset Management	DAM	The process of managing digital assets throughout their lifecycle. DAM systems manage, locate and retrieve specific digital content objects for use and reuse throughout diverse media. Commonly limited to traditional document formats. Vendors are quickly moving to include MAM.	Documentum FileNET Stellent
Content Management Systems	CMS	The strategy and technology for storing and indexing information from and about analogue or digital media.	Artesia
Website Content Management	WCM	Similar to CMS, focus is on website publishing workflow.	Interwoven Vignette
Media Asset Management	MAM	Similar to DAM, focus is on non-traditional document formats.	Convera Virage Cinebase
Enterprise Content Management	ECM	Similar to CMS, integrates into broader workflows.	Yahoo IBM Vignette
Digital Media Management Systems	DMMS	Same as MAM.	Same as MAM
Digital Content Distribution	DCD	The interaction of technologies, tools, and events involved in the circulation of text, sound, video, and data combinations over IP networks, between the points of content creation and the points of content consumption.	Akamai Digital Island Inktomi Speedera Virage

The above table represents a snapshot in time that will change rapidly over the next few years. These companies and their systems are entering a period of convergence where mergers and acquisitions will be common. Additionally, many of these technologies will broaden to include other areas, eventually converging to one or two areas. While any of the above mentioned technologies provide some form of video asset management, the focus of this section will be technologies that can capture, index and store video assets. The following table lists sample vendors of DAM systems and shows their functionality.

	Convera ScreeningRoom	Virage Video Logger
Video Capture Options	NTSC, PAL, SECAM, scene detection, text capture.	NTSC, PAL, SECAM, scene detection, text capture.
Output Formats	Windows Media, QuickTime, Real and MPEG formats, multiple bit rates.	AVI, Windows Media, QuickTime, RealVideo and MPEG formats, multiple bit rates.
Repository Options	RetrievalWare, ODBC compliant databases.	Virage Video Application Server, Oracle.
Search Options	Color/Shape/Texture search, Natural Language Search via RetrievalWare.	Via relational database or customized interface to commercial search engine.
Company URL	http://www.convera.com	http://www.virage.com

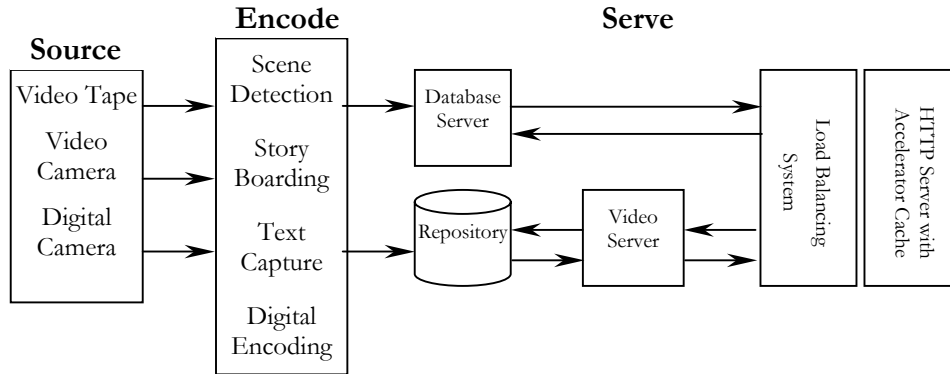
A great deal of IT planning is required to build an effective hardware and software solution for managing video assets. The size and number of video servers, amount of network bandwidth, size and speed of storage systems are all IT decisions that must be made. These decisions will require information such as: how many simultaneous users will be streaming video, how many simultaneous users will be searching the archive, what number of output formats will be offered, and what is the desired granularity of indexing. Clearly, these decisions can consume enormous amounts of money, not just initially, but over a sustained period. Many of these costs are re-occurring, such as network bandwidth, application software maintenance fees, leasing hardware, etc. As with all computer-related endeavors there are options. A pros/cons comparison of commercial, open source, in-house and ASP solutions is provided in the following table.

Solution	Pro	Con
Commercial DAM	Often highly automated, targeted to video content. Handles all standard formats. Powerful computer-aided indexing.	Expensive – initial and reoccurring costs. Specialized system – harder to sustain.
Commercial Content Management Systems	Generalized system – easier to leverage between groups, easier to sustain.	Not specific to video content. Very limited indexing capabilities.
Open Source	Much less expensive. Less feature rich.	Volunteer technology community may lose interest in supporting. Multiple Open Source components will be required to provide a “complete solution.”

Solution	Pro	Con
In-house	Just what you want/need.	Expensive to develop and sustain. Not as fast to put system together. Requires advanced skill set.
Application Service Provider	All issues handled by professionals. Unlimited growth potential. Often have integrated e-commerce solutions.	Re-occurring costs. Less control.

Putting It All Together—Schematic System Configuration

The following diagram is intended to give a high-level view of what may be required to provide a searchable video archive. Please note that such configurations must be planned for scalability, reliability and desired functionality.



Appendices

A) Bibliography & Further Resources

Introduction

1. Coalition for Networked Information. Retrieved March 14, 2003, from <http://www.cni.org>
2. Institute of Museum & Library Services (IMLS). Retrieved June 2, 2003, from <http://www.ims.gov/>
3. IMLS. (2001, November 6). *A Framework of Guidance for Building Good Digital Collections*. Retrieved June 2, 2003, from <http://www.ims.gov/scripts/text.cgi?/pubs/forumframework.htm>
4. Internet2. Retrieved June 2, 2003, from <http://www.internet2.edu>
5. National Initiative for Networked Cultural Heritage (NINCH). Retrieved June 2, 2003, from <http://www.ninch.org/>
6. NINCH. (2002, November). *Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials*. Retrieved June 2, 2003, from <http://www.nyu.edu/its/humanities/ninchguide/>
7. Performance Archive & Retrieval Working Group. Retrieved June 2, 2003, from <http://apps.internet2.edu/html/perfarchiveretrievalwg.html>
8. Performance Archive & Retrieval Working Group Charter, Retrieved June 2, 2003, from <http://apps.internet2.edu/html/perfarchiveretrievalwg-charter.html>
9. Performing Arts Data Service (PADS). Retrieved June 2, 2003, <http://www.pads.ahds.ac.uk/>
10. PADS. (2000). *Guide to Good Practice: Creating Digital Audio Resources*. http://www.pads.ahds.ac.uk:81/GGP_Audio (Old version, retrieved June 2, 2003. A reformatted version is available on the new PADS website; unfortunately under a dynamic URL. It can be reached from the home page under the 'advice' tab, in the 'Guides to Good Practice' category.)
11. PADS. (2000). *Guide to Good Practice: Creating Digital Performance Resources*. (Old version no longer available. The reformatted version is available on the new PADS website; unfortunately under a dynamic URL. It can be reached from the home page under the 'advice' tab, in the 'Guides to Good Practice' category.)
12. The Video Developer's Initiative (ViDe). Retrieved June 2, 2003, from <http://www.vide.net>
13. ViDe. (2002). *The Videoconferencing Cookbook Version 3.0*. Retrieved June 2, 2003, from <http://www.videnet.gatech.edu/cookbook/>

Preservation

1. National Library of New Zealand. Retrieved June 2, 2003, from <http://www.natlib.govt.nz/>

Intellectual Property

1. Amherst College. (2001, July). *Fair Use and Education Copying Guidelines for Teaching*. Retrieved June 2, 2003, from <http://www.amherst.edu/~deanfac/policies/copyingguidelines.html>
2. Ball State. (2002, March). *Fair Use and the Four Factors*. Retrieved June 2, 2003, from <http://www.bsu.edu/library/thelibraries/units/copyright/4fairuse.html>
3. *The Family Educational Rights and Privacy Act of 1974*. (FERPA). The U.S. Department of Education's FERPA index page is at: <http://www.ed.gov/offices/OM/fpco/ferpa/index.html> (Retrieved June 2, 2003). The *Legislative History of Major FERPA Provisions* (with further overview) is at: <http://www.ed.gov/policy/gen/guid/fpco/ferpa/leg-history.html> (Retrieved September 21, 2003).
4. Gemmill, Jill. (2002). *ViDe Resources: Consent/Release Form Samples*. Retrieved June 2, 2003, from <http://www.vide.net/resources/forms/>

5. Overing, Michael. (2001, June 5; modified 2002, April 4). Not So Public Figures. *University of Southern California Annenberg School of Communication Online Journalism Review*. Retrieved June 2, 2003, from <http://www.ojr.org/ojr/law/1017956813.php>
6. University of California San Francisco. (2003, March 10). *Paper/AV and Electronic Reserve Copyright Guidelines*. Retrieved June 2, 2003, from <http://www.library.ucsf.edu/edtech/reserve/cpguide.html>
7. University of Rochester. (1998, December). *Copyright Clearance and Fair Use at the University of Rochester*. Retrieved March 21, 2003, from <http://www.lib.rochester.edu/copyright/urpolice.htm>

Metadata

1. Baca, Murtha, ed. (1998). *Introduction to Metadata: Pathways to Digital Information*. Getty Research Institute. Retrieved June 2, 2003, from <http://www.getty.edu/research/institute/standards/intrometadata/>
2. Baca, Murtha. (1998). *Metadata Standards Crosswalks*. Getty Research Institute. A section of the above guide, retrieved June 2, 2003, from http://www.getty.edu/research/institute/standards/intrometadata/3_crosswalks/index.html
3. Consultative Committee for Space Data Systems. (2002, January). *Reference Model for an Open Archival Information System (OAIS)*. Retrieved June 2, 2003, from <http://www.ccsds.org/documents/650x0b1.pdf>. Also see NASA's Overview of ISO standardization efforts in this area (Retrieved June 2, 2003, from <http://ssdoo.gsfc.nasa.gov/nost/isoas/overview.html>) and the Research Libraries Group (RLG) OAIS overview, retrieved June 2, 2003, from <http://www.rlg.org/longterm/oais.html>
4. The Cover Pages. Retrieved June 2, 2003, from <http://xml.coverpages.org/>
5. Dublin Core Metadata Initiative (DCMI). Retrieved June 2, 2003, from <http://www.dublincore.org/>
6. Encoded Archival Description (EAD) Official EAD Version 2002 Web Site. [This site also provides some documentation for the now-superseded EAD Version 1.0.] Retrieved June 2, 2003, from <http://www.loc.gov/ead/>
7. Getty Museum. *Art and Architecture Thesaurus*. Retrieved June 2, 2003, from <http://www.getty.edu/research/tools/vocabulary/aat/>
8. Hunter, Jane. (2000, October). *Proposal for the Integration of Dublin Core and MPEG-7*. Retrieved June 2, 2003, from <http://www.egov.vic.gov.au/Documents/DCforMPeg7.doc>
9. Hunter, Jane, Martinez, Jose M., and Oltmans, Erik. (2000, July). *MPEG-7 Harmonization with Dublin Core: Current Status and Concerns*. Retrieved June 2, 2003, from <http://archive.dstc.edu.au/RDU/staff/jane-hunter/m6160.zip>
10. Instructional Management Systems (IMS). <http://www.imsproject.org/>. Especially relevant are the specifications for *Content Packaging*: <http://www.imsproject.org/content/packaging/index.cfm>, *Metadata*: <http://www.imsproject.org/metadata/index.cfm>, and *Digital Repositories*: <http://www.imsproject.org/digitalrepositories/index.cfm>. (All retrieved June 2, 2003.)
11. International Organization for Standardization (ISO). (2002, July). *MPEG-7 Multimedia Content Description Interface Overview (version 8)*. Retrieved June 2, 2003, from <http://ipsi.fraunhofer.de/delite/Projects/MPEG7/Documents.html>
12. International Organization for Standardization (ISO). (2002, October). *MPEG-21 Multimedia Framework Overview (version 5)*. Retrieved June 2, 2003, from <http://mpeg.telecomitalia.com/standards/mpeg-21/mpeg-21.htm>
13. Library of Congress, Cataloging Policy and Support Office. (2002). *Library of Congress Subject Headings Library of Congress, Cataloging Distribution Service*.

14. Machine Readable Cataloguing (MARC). Retrieved June 2, 2003, from <http://lcweb.loc.gov/marc/>
15. Metadata Encoding & Transmission Standard (METS). Retrieved June 2, 2003, from <http://www.loc.gov/standards/mets/>
16. Metadata Object Description Schema (MODS). Retrieved June 2, 2003, from <http://www.loc.gov/standards/mods/>
17. National Science Digital Library. (2002). *Metadata Primer Page* <http://metamanagement.comm.nsdlib.org/outline.html> and *Metadata Resources Page* (with crosswalks). <http://metamanagement.comm.nsdlib.org/IntroPage.html> (Both retrieved June 2, 2003.)
18. National Information Standards Organization (NISO). *Data Dictionary - Technical Metadata for Digital Still Images*. Retrieved June 2, 2003, from http://www.niso.org/committees/committee_au.html
19. OCLC/RLG Working Group on Preservation Metadata. (2002, June). *Preservation Metadata and the OAIS Information Model: A Metadata Framework to Support the Preservation of Digital Objects*. Retrieved June 2, 2003, from http://www.oclc.org/research/projects/pmwg/pm_framework.pdf
20. Shareable Content Object Reference Model Initiative (SCORM). (2002). *SCORM Overview*. Retrieved June 2, 2003, from <http://www.adlnet.org/index.cfm?flashplugin=1&fuseaction=scormabt> More information about SCORM's courseware/distance learning systems is also available on the XML Cover Pages, retrieved June 2, 2003, from <http://xml.coverpages.org/scorm.html>.
21. Text Encoding Initiative Consortium. (2001, August 15; revised 2003, January 6). *Welcome to the TEI Website*. Retrieved June 2, 2003, from <http://www.tei-c.org/>
22. University of California Berkeley. (2002, June 14). *GDM: A Descriptive Metadata Extension Schema for the Metadata Encoding and Transmission Standard*. Retrieved June 2, 2003, from <http://sunsite.berkeley.edu/mets/gdmSchema.html>
23. Visual Resources Association (VRA). (2002, February 20). *VRA Core Categories Version 3.0*. Retrieved June 2, 2003, from <http://www.vraweb.org/vracore3.htm>
24. Wactlar, Howard, and Michael G. Christel. (2002). *Digital Video Archives: Managing Through Metadata*. <http://www.informedia.cs.cmu.edu/documents/Wactlar-CLIR-final.pdf> (In: NDIIPP. (2002, April). *Building a National Strategy for Preservation: Issues in Digital Media Archiving*. <http://www.clir.org/pubs/reports/pub106/contents.html>) (Both retrieved June 2, 2003.)
25. World Wide Web Consortium (W3C). (2003, February 7). *Hyper Text Markup Language (HTML) Home Page*. Retrieved June 2, 2003, from <http://www.w3.org/MarkUp/>
26. W3C. Extensible Markup Language (XML). Retrieved June 2, 2003, from <http://www.w3.org/XML/>. Also see Robin Cover's Cover Pages for information on all sorts of SGML and XML languages, standards, and projects, also retrieved June 2, 2003, from <http://xml.coverpages.org/>
27. W3C. *Overview of SGML Resources*. Retrieved June 2, 2003, from <http://www.w3.org/MarkUp/SGML/>
28. W3C. *Resource Description Framework (RDF)*. Retrieved June 2, 2003, from <http://www.w3.org/RDF/>
29. W3C. *Synchronized Multimedia (SMIL homepage)*. Retrieved June 2, 2003, from <http://www.w3.org/AudioVideo/>

Digital Collection Software/Systems

1. Convera. Retrieved June 2, 2003, from <http://www.convera.com>

2. Virage. Retrieved June 2, 2003, from <http://www.virage.com>
3. Artesia. Retrieved June 2, 2003, from <http://www.artesiotech.com>

Further Resources

1. Berkeley Art Museum/Pacific Film Archive (BAMPFA). (2002). *Archiving the Avant Garde: Documenting and Preserving Variable Media Art*. Retrieved June 2, 2003, from http://www.bampfa.berkeley.edu/ciao/avant_garde.html
2. Council on Library and Information Resources (CLIR). (2001, May). *Folk Heritage Collections in Crisis*. Retrieved June 2, 2003, from <http://www.clir.org/pubs/reports/pub96/contents.html>
3. Johnson, Catherine J. and Allegra Fuller Snyder. (1999, July). *Securing Our Dance Heritage: Issues in the Documentation and Preservation of Dance*. CLIR. Retrieved June 2, 2003, from <http://www.clir.org/pubs/reports/pub84/pub84.html>
4. NDIIPP. (2002, April). *Building a National Strategy for Preservation: Issues in Digital Media Archiving*. CLIR. Retrieved June 2, 2003, from <http://www.clir.org/pubs/reports/pub106/contents.html>
5. Price, Laura, and Abby Smith. (2000, March). *Managing Cultural Assets from a Business Perspective*. CLIR. Retrieved March 21, 2003, from <http://www.clir.org/pubs/reports/pub90/contents.html>

B) Notes on Other Relevant Guides

This ‘current practice’ guide is but one of many guides and recommendations for digitizing cultural heritage materials. Among others that we used for background information and as a model for this document are guides from the National Initiative for a Networked Cultural Heritage (NINCH), the Performing Arts Data Service (PADS), the Video Development Initiative (ViDe), and the Institute for Museum and Library Services (IMLS) [Introduction 5, 9, 12, 2]. Of these, the broadest are NINCH’s *Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials* [Introduction 6] and “A Framework of Guidance for Building Good Digital Collections” [Introduction 3] from IMLS. They look at all aspects of building digital collections, with recommendations ranging on levels from overall project management to specific practices, and they deal with the digitization of all sorts of original media. Both guides are in-depth resources, with links to further information.

More specialized are two guides from PADS in their “Guide to Good Practice” series: *Creating Digital Performance Resources* [Introduction 11] and *Creating Digital Audio Resources* [Introduction 10]. They focus specifically on the process for live media, and do not, for example, go into the digitization of text or still images. They are both created with a United Kingdom audience in mind, and as such deal with copyright and other intellectual property issues somewhat differently than this document.

ViDe’s *Videoconferencing Cookbook* [Introduction 13] focuses on videoconferencing applications, hardware, and software. It was interesting to us as a model, especially for its form and for its section on broadcasting and archiving videoconferences via streaming video.

C) Authorship and Contact Information

Version 0.9 of this document was authored by the Performance Archive and Retrieval Working Group, jointly sponsored by Internet2 and the Coalition for Networked Information. The Working Group will be integrating additional comments in the future and releasing version 1.0 of the document as a result. Comments on this document should be directed to Ann Doyle, Working Group Co-Chair and Internet2 Program Manager for Arts and Humanities Initiatives at adoyle@internet2.edu

These guides to good practice are the product of the joint Internet2/CNI Performance Archive and Retrieval Working Group. Working Group members have made major and varied contributions to the guides, including but not limited to editing, reviewing, and drafting sections, as well as proposing initial outlines and topics for consideration and contributing general subject expertise.

Working group members who have made specific significant contributions to individual sections are indicated in the list below.

Working Group Members

Richard Albertson, Star Valley Solutions	Carell Jackimiek, University of Manitoba
Melissa Beck, University of Michigan	Jim Kerkhoff, University of Texas at Austin
Robert Bender, University of Missouri – Columbia	Brian Lavoie, OCLC - Online Computer Library Center
Johannes Birringer, Ohio State University	* Joan Lippincott, Coalition for Networked Information — Sponsor (<i>Institutional Policies</i>)
Pamela Bloom, New York University	Judy Lubben, University of Illinois Urbana-Champaign
Tom Bramel, Library of Congress	Clifford Lynch, Coalition for Networked Information — Sponsor
* Jim Brei, Proquest Information and Learning (<i>Search and Retrieval, Digital Collection & Asset Management Systems</i>)	John Mitchell, Arizona State University
Markus Buchhorn, Australian National University	Sharon Moskwiak, Internet2
Beth Chancellor, University of Missouri – Columbia	James Oliverio, University of Florida
Ann Doyle, Internet2 — Sponsor and Chair	* Bill Savage, Proquest Information and Learning (<i>Institutional Policies</i>)
Jim Evans, Rensselaer Polytechnic Institute	James Schaefer, Georgetown University
* Amy Friedlander, Council on Library and Information Resources, detailed to the Library of Congress (<i>Preservation</i>)	Peter Sparling, University of Michigan
* Philip Galanter, New York University (<i>Audio Capture</i>)	Art St. George, University of New Mexico
Belle Gironde, State University of New York, Albany	Henry Szujewski, University of Illinois at Urbana-Champaign
Alan Green, Ohio State University	Susan Topol, Internet2
David Green, NINCH - National Initiative for a Networked Cultural Heritage	Stephen White, University of Wisconsin, Milwaukee
* John Grozik, University of Wisconsin, Milwaukee (<i>Video Capture</i>)	Lenita Williamson, University of Maryland
Ted Hanss, Internet2	Rembrandt Wolpert, University of Arkansas
Leslie Hitch, Northeastern University	

Mike Holcomb, University of Arizona
Lorna Hughes, New York University

Thanks and appreciation also need to go to the following people outside the Working Group:

- * Martha Anderson, who co-wrote the Preservation section with Amy Friedlander.
- * Grace Agnew and Mairead Martin, whose papers greatly contributed to the Digital Rights Management portion of the Intellectual Property section.
- * Jill Gemmill, who collected and provided examples of performer/participant release forms.
- * Mick Goulash and Daniel LaFavers, who assisted in writing the Search & Retrieval and the Digital Collection & Asset Management Systems sections.
- * Ted Hertzog, whose initial motivation spurred the forming of the Working Group, and who was its first co-chair with Ann Doyle.

Finally, the Working Group intern, Leilani Dawson, compiled and edited the guides overall, and drafted some sections based on the Working Group's direction and input.