NATIONAL ENDOWMENT FOR THE HUMANITIES

SAMPLE APPLICATION NARRATIVE

Preservation Education and Training Grants
Institution: Cornell University Library Department of Preservation and Conservation and the Cornell Institution for Digital Collections

Project Summary

The Cornell University Library Department of Preservation and Conservation and the Cornell Institute for Digital Collections request funding from the National Endowment for the Humanities to support an education and training program, ***Digital Preservation Management: Effective Short-Term Strategies for Long-Term Problems.*** By incorporating a distance learning tool and an intensive, limited enrollment seminar, this innovative program will employ a tiered approach to effective decision making in an age of technical uncertainty. It will build on Cornell's experience in offering several series of workshops over the past decade, in creating Web-based training tutorials, and in conducting applied research in digital preservation. The educational offerings will present an integrated approach to developing digital preservation programs that incorporate technical, financial, organizational, and policy aspects encompassing the full life cycle of digital objects. These programs will stress short-term risk reduction strategies while research and development goes forward in creating longer-term solutions. The full range of digital content created, collected, served by—or linked to—cultural institutions will be addressed. Issues to be covered include: program planning, management and evaluation; risk assessment; cost benefit analysis; legal issues; file formats, standards, metadata and their connection to preservation implementation; storage and maintenance; disaster planning; the relationship between preservation and access; preservation strategies; and technology forecasting.

Cornell seeks NEH funding to undertake the following work:
- Develop and maintain a Web-based, self-directed tutorial covering the basics of digital preservation for cultural repositories.
- Offer 8 one-week workshops with a limited enrollment of 24 participants each.
- Provide partial tuition support for the 192 participants attending those workshops.

This program is intended for those contemplating or implementing digital preservation programs in libraries, archives, and other cultural institutions. The goals of this initiative are to foster critical thinking in a technological realm and to provide the means for exercising practical and responsible stewardship of digital assets in a time of uncertainty and flux. Cornell's curriculum will incorporate a pedagogy that ensures ready access to basic technical information, is well suited to adult learning styles, and takes best advantage of a limited enrollment seminar. Four different audiences will be targeted for 2 workshops each, including research libraries; colleges and mid-sized universities; archives and historical societies; and small, underserved cultural institutions, such as the historically black colleges and universities.

Introductory information will be provided via a stand-alone tutorial. This Web-based tutorial will serve as a prerequisite to participation in the Cornell workshop to ensure a more uniform understanding among attendees, and will be offered in several languages (e.g., Spanish, French, Portuguese). The workshop will be an advanced, results-driven educational offering in which small groups work together to define and address a continuum of issues raised by the challenges associated with preserving digital objects.

Table of Contents

VII.   APPENDIXES

Appendix A: Tutorial Curriculum

Appendix B: Workshop Curriculum

Appendix C: Table of Digital Technology Workshops

Appendix D: Testimonials

Appendix E: Gantt Chart

Appendix F: Checklist of Workshop Responsibilities

Appendix G: Evaluation and Survey Forms

Appendix H: Resumes and Biographical Statements

Appendix I: Keynote Speakers

Appendix J: Advisory Committee

Appendix K: Letters of Support

I.   PROJECT SIGNIFICANCE

The Challenges of Digital Preservation

In recent years, the national media has treated the American public to horror stories regarding the fate of important documents available only in digital form. Oft-repeated tales chronicle the loss of valuable scientific and business data stored on punch cards or 8-inch floppy disks or written in computer languages long dead. We read of the loss of census data and land use records, POW and MIA files from the Vietnam War, 1970's satellite photos documenting rain forest destruction along the Amazon basin, or data collected during NASA's 1976 Viking mission to Mars. After Germany was reunified in 1990, many East German data files were no longer legible and necessary documentation was incomplete or missing. Estimates put the average life expectancy of a Web page at somewhere between 75 days and 2 years, and of those that survive, nearly all undergo some change in content within a year.   In late 2000, George W. Bush's transition team instructed all federal agencies to remove information from their Websites specifically related to the Clinton administration. The Berkeley Digital Library SunSITE periodically releases a list of electronic government publications no longer available on the Internet. Thus far, no provision has been made to ensure that these publications will be available to future users, and no explanation is given as to why they have been discontinued.[1]

The fragility of digital media is also worrying. Experts predict that current media types will begin to decay in one to ten years, depending on the quality of the tape or disc and its storage conditions. Every two-to-three years sees the introduction of a new generation of hardware or software, with backward compatibility most often guaranteed for only a limited time thereafter. Migrating data to new technologies is analogous to the continuous painting of the Golden Gate Bridge—when one gets to the end, the process starts over immediately. Migration also comes at a price, is fraught with legal barriers, and has its own set of risks. The FDA has reported data loss by pharmaceutical firms migrating drug-testing data and that blood pressure numbers were "randomly off" by up to eight digits from those in original reports.

These rather dramatic tales highlight some of the major dangers associated with managing digital content over time. The goal of digital preservation is to maintain the ability to display, retrieve, and use digital material in the face of rapidly changing technological and organizational infrastructures. Unfortunately, there is no single best way to do just that, nor is there agreement on long-term solutions. Even in the short-term, librarians and archivists must understand their own institutional requirements before they can begin to identify which combination of policies, strategies, and tactics are likely to be most effective in meeting their needs.

One of the main challenges in digital preservation is the pace of technological change, which according to Raymond Kurzweil, founder of several technology-based companies, doubles every decade.[2] Although such advances promise more effective solutions for information-management problems, they also set a moving target that hinders long-term planning. In order to assess and manage risks associated with digital preservation, institutions need policies that reflect today's knowledge and capabilities, as well as the fluid nature of technology itself.

---

[1] Getty Conservation Institute and Getty Information Institute, "Time and Bits: Managing Digital Continuity," www.longnow.com/10klibrary/TimeBitsDisc/press.html; Michael Wettengel, "Unification and E-records: The example of East Germany's 'Kaderdatenspeicher,'" http://www.caldeson.com/RIMOS/kader.html; Jeffrey Benner, "Oil and Websites Don't Mix," www.wired.com/news/politics/0,1283,42536,00.html; "Electronic information no longer available on the Internet: May 10, 2000 " sunsite.berkeley.edu/GODORT/2k0508missing.html
[2] As cited in Denise Troll, "How and Why are Libraries Changing?" Council on Library and Information Resources, 2001.

A second major challenge is the increasing dependency of libraries and archives—and their users—on digital content and Web-accessible resources and services. In the past decade, the Internet has grown exponentially in size and complexity. A much-publicized study conducted at the University of California at Berkeley last year estimated that "the world's yearly production of print, film, optical, and magnetic content would require roughly 1.5 billion gigabytes of storage," with 93% of the information produced in digital form.[3] In 1993, an estimated 10% of U.S. libraries provided their users with Internet access: today nearly all of them do.[4] Research suggests, however, that more users of cultural electronic resources are conducting their work outside the walls of libraries and archives. In some institutions, 75% or more of the collection use occurs off premises.[5]

The Need for Education and Training
Recent surveys indicate that institutions cannot afford to wait for the ideal digital preservation solution to come along. From a technological perspective, digital collections are recent creations, using immature technologies that carry distinct risks to their long-term viability. In 1998, Margaret Hedstrom and Sheon Montgomery conducted a survey on behalf of the Research Libraries Group on "Digital Preservation Needs and Requirements in RLG Member Institutions."[6] Of the 54 institutions that completed the survey, most reported holding collections that were then one-to-three years old. By this reckoning, a large number of digital files held by libraries are potentially in danger from technology obsolescence. Equally disturbing is the fact that two-thirds (36 of 54) of respondents did not have written policies for digital preservation and only eight institutions considered their staff as "expert" in digital library activities, while forty-four ranked their knowledge at the "novice" or "intermediate" levels. Among the 30 research libraries included in the RLG survey, two-thirds reported institutional responsibility for preserving digital materials, whether or not they had formal preservation policies and trained personnel. In assessing this information, Margaret Hedstrom spoke of the lag between the development and implementation of best practices, and the gap between current guidance on digital preservation and institutional capacities to follow through. She recommended the following priorities:
- Implement the best practices that already exist
- Provide training
- Share successful policies, practices, and standards
- Focus research efforts on emerging promising solutions

A second survey, conducted by the Preservation Committee of the Association of Research Libraries in the summer of 2000, identified the seven most common preservation challenges facing ARL libraries, with 46% of respondents ranking funding as their top concern. The next greatest concern identified was digital preservation (36%).[7]

These findings are consistent with a -survey conducted by the Digital Library Federation (DLF) in early 2001. Fourteen out of twenty-one institutions reported having no formal preservation policy, although over half reported having digital preservation responsibilities for commercial data under lease and sixteen of twenty-one for library-produced digital materials.   Where preservation is concerned, only the library-

---

[3] How Much Information Project, Peter Lyman, Hal R. Varian et.al, http://www.sims.berkeley.edu/research/projects/how-much-info/
[4] Developing an Agenda for Libraries and Information Policy Issues -- Introduction. The following is excerpted from a column by OITP Director Rick Weingarten available in ALA's online TechSource magazine. http://www.ala.org/oitp/prinintro.html (May 2001?)
[5] Denise Troll, op. cit., p. 11
[6] www.rlg.org/preserv/digpres.html
[7] Unpublished summary of the ARL survey.

produced holdings command real attention. Other kinds of digital content (e.g., e-journals, e-prints, university records) are most at risk.

It has been seven years since the International Council on Archives conducted a survey of electronic records programs, but all indicators are that while digital preservation research has increased significantly in the interim, the number of implemented institutional programs remains small.[8] Similarly, the National Historical Publications and Records Commission of the National Archives, has funded electronic records research since the early 1990s. To date, these projects have not resulted in a significant increase in implemented programs.

Consider the following situations:

Institution A relies on a commercial supplier of electronic resources and struggles with what to do with their physical counterparts and concerns over the long-term accessibility of the digital content.

Institution B is interested in digitizing sound recordings but is uncertain about the connection between content creation and long-term viability.

Institution C holds material in an obsolete format, and must determine the hows and whys of legacy systems and data migration.

Institution D is interested in providing reliable access to Web-based resources, but wonders about the options for exercising remote control.

Institution E must establish a central depository for digital assets, and Institution F is considering outsourcing digital preservation services: both seek advice on identifying minimal requirements.

Institution G must come to terms with the archival responsibility for its electronic records and faculty Web sites, and doesn't know where to begin.

These real-life concerns speak to the need for an education and training program that prepares cultural institutions to make informed decisions about the viability of digital content. The impulse to embrace things digital is strong, but too often insufficient thought is given to infrastructure requirements—costs, personnel, systems, and policies—and reality falls short of the promise. Information professionals can little afford to make mistakes in developing programs to sustain digital assets. They must assess the pros and cons of technology choices in a cultural context.

Why this Educational Program?
The time is right to undertake an intensive weeklong workshop devoted to developing short- term digital preservation strategies for two reasons: (1) our understanding of the issues has matured sufficiently to offer practical advice and provide the means for informed decision making, and (2) cultural institutions are increasingly relying on digital assets and assembling an array of at-risk resources. Digital content stewards must begin to address digital preservation responsibilities now.

This program is intended for those who are implementing or have been charged with developing digital preservation programs in libraries, archives, and other cultural institutions. Much work is underway within the United States and abroad on devising long-term digital preservation strategies. It is clear from these efforts that to date no single strategy has emerged that is suited to all formats, all institutions, and all content. What is needed is a practical, intensive workshop designed to assist participants in developing realizable and effective short-term digital preservation programs based on workable solutions.

---

[8] See "Electronic Records Programs: Report on the 1994/1995 Survey," ICA Study 9, 1996.

Information best practices to manage digital preservation programs in the short-term must be implemented even as we await longer-term solutions. Cliff Lynch recently noted, "We know that we do not fully understand how to preserve digital content; today there is no 'general theory,' only techniques."[9]

Cornell brings several distinctive perspectives to the proposed training program:

- A management bias will drive the curriculum. Organizational commitment is critical to successful program implementation. The objective will be to provide the tools and support to bring participants responsible for digital assets to the brink of implementation and to assess programmatic alternatives in light of their own needs.
- Increasingly, digital preservation initiatives are interdisciplinary and remove traditional lines between libraries, archives, museums and other cultural institutions. The workshop will build on the convergence of professional and organizational concerns to provide the most relevant information on standards, best practices, research initiatives, and model implementations that offer the greatest potential for short-term planning and interim solutions. The trainers and the participants will represent a range of domains, including information technologists as well as cultural professionals.
- Although there are many ongoing digital preservation research projects, there are too few implemented preservation programs. Organizations must define the type, size, and scope of program to suit their needs and resources. The workshop will offer practical approaches, progressive development, and incremental implementation based on current research findings, best practices, and relevant standards.
- Keeping up in this rapidly changing technical environment is one of the greatest challenges to trainers and practitioners. The modular structure of the workshop and tutorial will support the regular updates and additions that will be required in the course of this three-year endeavor.

Other Informational and Educational Offerings
The Cornell training program will offer vetted guidance on how to deal in the short run with significant preservation challenges. Certainly there is a plethora of available information on the issues and challenges. The National Library of Australia maintains a comprehensive subject gateway to digital preservation resources, called PADI (Preserving Access to Digital Information), which includes current information on events, organizations, policies, strategies, and guidelines. There are other well-respected Web sites that offer valuable information on digital preservation, but none provides an online resource similar to the tutorial proposed here.[10] In recent years conferences have focused on digital preservation, such as the Preservation 2000 conference held in York, England last December, and the OCLC/Preservation Resources' symposium on "Digital Past, Digital Future: An Introduction to Digital Preservation, " held at the ALA annual meeting in June of 2001. Appendix C offers a comparative table of workshops that touch upon digital preservation, including those offered by regional organizations. Like the current Cornell workshop (see *Teaching Experience)*, most of these focus on digital imaging—a subset of the larger universe of digital content that includes electronic records, text files, sound recordings, moving images, geo-spatial data, and dynamic datasets. Letters of support from the organizers of several of these workshops attest to the value of an intensive weeklong workshop devoted exclusively to digital preservation. One workshop leader wrote "I can unreservedly attest to the quality of instruction, currency of content, and genuine connectedness with participants that are your hallmark." (See Appendix K.)

---

[9] Clifford Lynch, "The Battle to Define the Future of the Book in the Digital World, *First Monday* (June 2001)
[10] Sites surveyed include PADI, http://www.nla.gov.au/padi/, JISC http://www.jisc.ac.uk/; eLib, http://www.ukoln.ac.uk/services/elib/, AHDS, http://ahds.ac.uk/viable.htm, UKOLN, http://www.ukoln.ac.uk/

Response to the Proposed Workshop
To gauge interest in this workshop, an inquiry was submitted to the listserv of the Preservation Administrators Discussion Group (PADG) of ALA. We received forty-two positive responses within twenty-four hours. This is a high response rate from a listserv of 337 subscribers. We also surveyed participants in the 2000 workshop series at Cornell (See Appendix G for the survey questionnaire). Of those who had already attended an intensive workshop at Cornell, 90% indicated that they would be very interested in or would consider attending such a workshop. Appendix D contains sample comments received from these two informal surveys. The need for intensive training is further evidenced by the continuing demand for Cornell's current digital imaging workshops. We do little advertising, yet every workshop has been oversubscribed. The July and October 2001 workshops filled within two hours of the registration opening.

On May 17, 2001, the preservation administrators of the eleven Comprehensive Research Libraries of New York (the "Big 11") met in Albany, NY.   At the meeting, Cornell representatives discussed the possibility of conducting a digital preservation workshop series. Attendees expressed unanimous support, and it was agreed that the Big 11 would serve as the beta test group. Cornell will submit a grant to the New York State Program for the Conservation and Preservation of Library Research Materials in January 2002 to support the Big 11 expenses in attending this first workshop in exchange for their participation in a formal evaluation of the curriculum and accompanying tutorial.

Who Will Benefit From This Program?
Our objectives for the new education and training program are to:
- Develop a practical skills set required to implement short-term digital preservation programs based on national standards, accepted theory and practice, and a common vocabulary.
- Sharpen critical thinking and priority setting capabilities, especially in the technical realm.
- Foster goal-driven, not technology-driven, digital preservation programs to empower non-technical experts in selecting and using the right tools to accomplish their work.
- Build group problem solving skills by sharing ideas and expertise among the participants.
- Explore thematic variations by tailoring the curriculum for four different audiences.

Introductory, base level information will be provided via the self-directed tutorial, which will be designed as a stand-alone, Web-based tool. Intended to serve a broad national and international audience, the tutorial will be offered in several languages. Its availability will improve awareness across the spectrum of institutions and individuals involved in the creation and dissemination of digital content. The tutorial will also serve as a prerequisite to participation in the Cornell workshops, ensuring a more uniform level of understanding among attendees.

The workshop will be an advanced, results-driven educational offering in which small groups work together to define and address a continuum of issues raised by the challenges associated with preserving digital assets. Four key audiences will be targeted with two workshops each. The first target group is major research libraries, where the tendency to create independent capabilities still prevails. The workshops aimed at this group will focus on the establishment of full archival repositories, but also emphasize the need to embrace collaborative approaches. The second group, college and mid-sized universities, may be quicker to appreciate collaborative solutions, and in these workshops, we will focus on such arrangements as well as defining minimal requirements at the local level. Archives, historical societies, and special collections represent distinct constituencies, and the particular requirements for institutional records and Web-based content management will be stressed. The needs of the final targeted group, underserved institutions, will be addressed by stressing basic, economical preservation management practices and the pros (and cons) of investing in digital content development. We will also

solicit the participation of policy makers, funding officers, and congressional staffers who must become familiar with the dimensions of the digital preservation problem. Indirectly, scholars and other researchers will benefit from continuing access to research collections made accessible or only available in digital form.

The workshop will provide participants with the means to move beyond theoretical constructs to implementation strategies grounded in practical solutions emphasizing short-term programs for managing risk as research and development continues in assessing long-term solutions. Participants will be encouraged to evaluate how well current/emerging standards and best practices relate to their own institutional needs. Consider the following. Much work has gone into identifying metadata to support long-term preservation management, but what evidence exists that collecting such metadata enables preservation action or good decision-making? In the past eight years, we've learned about the ephemeral nature of Uniform Resource Locators (URLs), and that there are a number of initiatives to implement Uniform Resource Names (URNs). Which competing scheme is the most appropriate based on particular needs (PURLS, Handles, DOIs), and what are the necessary steps to take—and the institutional requirements—to establish such a service? There are strong advocates of migration strategies and equally strong voices arguing for emulation-based approaches, but how does one go about evaluating the relevance of each approach based on the specifics of the institution, the digital content, user requirements, and available resources? Object-based approaches are being investigated to support the integration of required metadata with the relevant digital content to be preserved. These approaches may also be used in conjunction with migration and emulation approaches. What does object-based preservation look like and what is the potential advantage of this approach? The workshop will provide participants with the means to evaluate the pros and cons associated with various organizational and technological choices. It will also serve as a living laboratory to accelerate the diffusion of digital preservation theory into the pragmatic but experimental realm of different institutions facing a variety of preservation challenges. Though its graduates, the workshop will lead to additional test-beds for various approaches, help weed out the impractical ones, and refine the ones with the most promise.

II.   HISTORY

The Cornell University Library is eminently qualified to undertake this work. It has earned an international reputation for technology leadership and training. This year, the Department of Preservation was honored with the LITA/Library Hi Tech Award for Outstanding Communication for Continuing Education in Library and Information Science for its "contributions in research, continuing education and information sharing to increase the ability of libraries to preserve information for future generations."

The Preservation Department pioneers in the development of theoretical constructs, such as the establishment of quality benchmarks for digital imaging, risk management, and in advocating a managerial approach to the consideration and implementation of digital programs. In the early 1990s, the staff conducted groundbreaking research into digital conversion requirements for books and journals, and their recommendations were subsequently adopted by major research libraries in the U.S. and abroad as well as by such important initiatives as JSTOR. Additional research has focused on conversion requirements for other source documents, on image quality and user requirements in presentation, and for the past five years, on digital preservation. Through an array of products and services—publications, workshops, management tools, research reports, and an online tutorial—the staff has presented this research in continuing education resources that are approachable, authoritative, and of practical utility.

Teaching Experience
Until Cornell began to offer weeklong workshops in 1994, there was no comprehensive continuing education program devoted exclusively to digital imaging. Since then Cornell has offered twenty-four workshops in Ithaca, New York that have been attended by over 700 individuals. The National

Endowment for the Humanities supported the most recent digital imaging workshop series (1999-2001) entitled "Moving Theory into Practice" (MTIP), which was designed to assist participants in the transition from project-oriented efforts to on-going programs.   In addition, from 1998-1999, the department taught a two and a half-day workshop on Managing Digital Imaging Projects (MDIP) for the Research Libraries Group. The focus of the workshop was to investigate managerial concerns in developing and implementing digital imaging projects. RLG offered the workshop eight times at various locations in the US, the UK, Australia, and Canada. The 1996 New York State guidelines for the Library Services and Technology Act made specific mention of the Cornell workshops as an eligible expense for applications to the "Electronic Digital Libraries Initiatives" program.

Participants' evaluations in these workshops have been consistently high. On average, they ranked them at 3.8 on a scale of 1-4, with 4 representing excellent. Of particular significance, the confidence level of participants in applying what they learned in the workshops rose steadily through the years: from 1994-1998 the confidence level averaged 3.44 out of 4; in the MDIP workshops it increased to 3.6, and it reached 3.7 in the 1991-2001 workshops. Indicating a shift in emphasis away from technology and toward managerial concerns, participants ranked as most useful the sessions focusing on management, mainstreaming, and digital preservation strategies. Comments from participants confirmed the usefulness and portability of the Cornell approach to other institutions. One participant put it "It is up to those of us who attended the training to exert our influence on our host institutions to do things right. You certainly have given us the tools to do so." See additional comments in Appendix D.

Outreach
In October 2000, the Department released the first online digital imaging tutorial that provided up-to-date and interactive content made freely available to a global audience. In response to user demand, staff added a Spanish version in March 2001 and a printable PDF version in May 2001. Participants in the last five workshops rated the tutorial at 3.65, on a scale of 1 to 4.
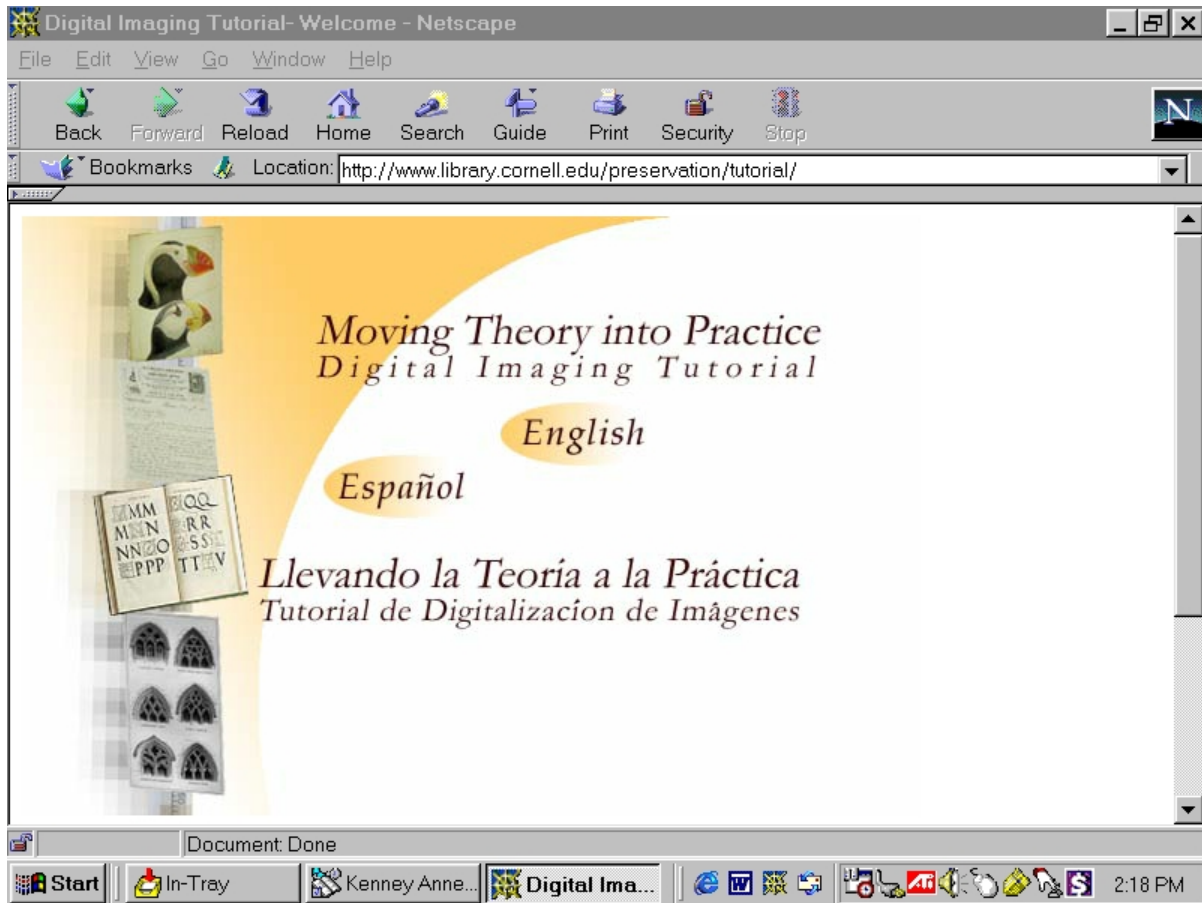
Figure 1. Home Page to Cornell's Digital Imaging Tutorial
www.library.cornell.edu/preservation/tutorial/

Response from other users has been strong and positive. Shortly after its public release, *The Scout Report*, published by the Computer Sciences Department, University of Wisconsin-Madison, highlighted the tutorial in the November 3, 2000 issue, indicating that it "provides a comprehensive overview of the issues involved in translating photos and documents into electronic form. This online tutorial explains the various steps, the technology behind them and some of the organizational challenges that they present. [It] provide[s] easily-accessible information that should be of increasing interest to anyone involved with digital publication and archiving of existing works."

The number of daily tutorial users averages one hundred—one third of whom are accessing the Spanish version. Since the printable version was made available, the number of successful requests for the PDF file has averaged 200/day for the English version and 21/day for the Spanish version. Of these users, 5-10/week submit reference queries to the staff, the majority from Spanish speakers (see Appendix D for sample letter and response).

A second indicator of the tutorial's success is the interest expressed in providing versions in other languages. The Programa Sociedade da Informação, Ministério da Ciência e Tecnologia in Brazil has proposed underwriting a Portuguese translation; the UNESCO Memory of the World Program and the Secretary General of the International Council on Archives have expressed interest in translating the tutorial into French (one of the two official languages of the ICA). The Social Sciences Research Council

supported the production of 50 CD copies of the tutorial for participants in a three-day preservation-planning meeting of Cuban librarians and archivists, held in Havana, June 14-18, 2001.

The tutorial has already been used in several graduate level courses in information science. Dr. Helen Tibbo, School of Information and Library Science, University of North Carolina at Chapel Hill, used it as a principal resource in her Fall 2000 Information Technologies for Cultural Heritage Information Professionals course (see letter of support). Professor Gretchen Whitney of the School of Information Sciences at the University of Tennessee included the tutorial in her Spring 2001 Information Networks Applications course, citing its "stunning navigation techniques." Elizabeth Shaw also used the tutorial in her Spring 2001Digital Preservation course at the University of Pittsburgh School of Information Sciences.

Publications
In 1996, Cornell University Library published *Digital Imaging for Libraries and Archives,* by Anne R. Kenney and Stephen Chapman, the first definitive guide to imaging technologies and processes designed expressly for the library and archival community. This book won the Society of American Archivists prestigious Leland Prize for "writing of superior excellence and usefulness." One reviewer on the *Archives & Archivists* list recommended it as "the most useful $75.00 we have spent at our library in a long time." Over 2,500 copies have been sold.

In 2000, the Research Libraries Group published the second monograph produced by the Preservation Department, *Moving Theory into Practice: Digital Imaging for Libraries and Archives*, co-authored by Anne R. Kenney and Oya Y. Rieger. This volume advocates an integrated approach to digital imaging programs, from selection to preservation, with a heavy emphasis on the intersection of institutional, cultural objectives and practical digital applications. SAA recently chose this book to receive the 2001 Leland Prize (representing only the fourth time that an author has won this prize twice since its inception in 1959). In the first year of its availability, nearly 2,000 copies have been sold.

Information changes rapidly in an emerging field. To meet the demand for up-to-date news and analysis, the Department produces and edits a bimonthly electronic newsletter for the Research Libraries Group, entitled *RLG DigiNews* (www.rlg.org/preserv/diginews/). Through its editorship, the staff brings timely information to practitioners struggling with the technical, legal, and managerial aspects of digital imaging and digital preservation. Beginning with the February 2000 issue, the coverage of digital preservation issues was significantly increased and specifically highlighted by a new logo, ⊙⊗ⓟ, which RLG may trademark to denote all resources they commit to maintain. Digital preservation coverage for the past year included feature articles (7), editor's interviews (2), FAQs (1), Highlighted Web sites (1), and RLG News (2). RLG reports that this publication has become one of its most popular electronic resources, with logins approaching 80,000 last year, representing readers from over 25 countries.

Current Digital Preservation Research
Five years ago, the Department of Preservation began to focus its research away from digital imaging to issues associated with digital preservation. This shift in emphasis was motivated by concerns over the long-term viability of Cornell's digital image collections and also by a growing awareness that while digital preservation is one of the major challenges facing cultural institutions, little applied research has been conducted. The Department investigates approaches that can directly benefit Cornell and other cultural institutions. The following projects have helped prepare the staff to undertake this new training program.

Risk Management of Digital Information: A File Format Investigation
www.clir.org/pubs/abstract/pub93abst.html

The report of this project, published by the Council on Library and Information Resources, details one of the only systematic studies devoted to assessing the risks to digital file formats during migration. The project assessed risks in general and those specifically associated with migrating digital images and numeric file and identified many of the vulnerabilities associated with different features of digital objects, including data integrity, metadata, and functionality. The report includes a risk assessment workbook that will help librarians and archivists identify potential risks involved in file-format migration.

Preserving Cornell's Digital Image Collections: Implementing an Archival Strategy, 1999-2001
www.library.cornell.edu/imls/
The goal of this project, funded by the Institute for Museum and Library Services was to plan and implement a preservation strategy for Cornell Library's digital image collections. Accomplishments include an inventory of the collections created over the past decade, an investigation of current and emerging file formats for long-term utility, a study of functional requirements for storage, and draft recommendations for preservation metadata. This project also assessed resource needs in terms of staff, equipment, space, time, and finances for a ten-year maintenance program. The major accomplishment was to develop policies and procedures that form the foundation of a centralized digital repository for Cornell University Library. The recommendations for digital image creation and metadata served to define requirements for the Digital Library Federation's proposed digital preservation master files to be included in a DLF registry, www.clir.org/diglib/collections/reg/regsum.htm.

Project PRISM (Preservation, Reliability, Interoperability, Security, Metadata), 1999-2003
www.prism.cornell.edu/
This DLI2 project is a four-year collaborative effort between the library and Cornell's Computer Science Department to investigate and develop policies and mechanisms needed for information integrity in the context of a component-based digital library architecture. The key research areas include long-term survivability of digital information, reliability of information resources and services, interoperability, security, and metadata that makes it possible to ensure information integrity in digital libraries.

Project Harvest, www.library.cornell.edu/harvest/
In December 2000, The Andrew W. Mellon Foundation launched the Mellon Electronic Journal Archiving Program, a one-year effort to plan the development of digital archives for long-term retention of scholarly journals in digital form. Cornell's contribution is Project Harvest, which focuses on agriculture e-journals. Cornell has taken a leading role in preserving and providing access to agriculture literature, in part through its work with USAIN (United States Agricultural Information Network), which oversees the National Preservation Program for Agricultural Literature. Through negotiations with publishers, the library will create a model agreement for the deposit of scholarly journals in agriculture and other scientific disciplines. Planning will be carried out on the technical design for a repository, and an organizational model will be developed for managing electronic resources over time. This focus will also enable the library to investigate issues related to scholarly acceptance of an electronic journal archives.

Cornell Institute for Digital Collections, cidc.library.cornell.edu/
CIDC is a cross-disciplinary team co-directed by Peter Hirtle and Anne R. Kenney established to explore the use of emerging technologies in providing greater access to cultural and scientific collections. CIDC has made many of Cornell 's rare and fragile collections available through the Web. Partnerships with cultural institutions and corporations have permitted CIDC to create digital collections from around the world. CIDC also conducts applied research to test and evaluate the utility of such resources for the Cornell community as well as diverse global audiences. During the past three years, CIDC's path-breaking projects have run the gamut from high-resolution production museum digital photography, to metadata standards for the performing arts, to the use of XML for EAD finding aids. CIDC staff have

lectured and consulted on digital capture, storage, project management, user interfaces, and rights management issues.

III.   METHODOLOGY AND STANDARDS

Curriculum Content
Appendixes A and B provide detailed outlines for the workshop and the tutorial. The tutorial will focus on vocabulary and key concepts and basic organizational and technical infrastructure requirements. It will include sample files, simulations, links to other resources, hands-on exercises, computational programs, and the ability to send comments and questions directly to Cornell staff. The workshop will be based on practical management experience, on the research conducted here and elsewhere, and on the expertise of outside speakers to address key areas. The workshop will not push one-size-fits-all solutions but encourage participants to think through choices and consequences. The curriculum focuses on five areas: institutional readiness and assets assessment; program components and models; digital assets in-depth— examples and scenarios; preservation approaches and techniques; and creating an institutional digital preservation framework.

The curriculum will advocate a holistic approach to preservation management. Participants will gain an understanding of the organic nature of digital preservation, with interdependencies connecting goals, resources, and processes involved in managing digital assets throughout their full life cycle – from initial selection to providing continuing access in the face of technology obsolescence, as illustrated in Figure 1.
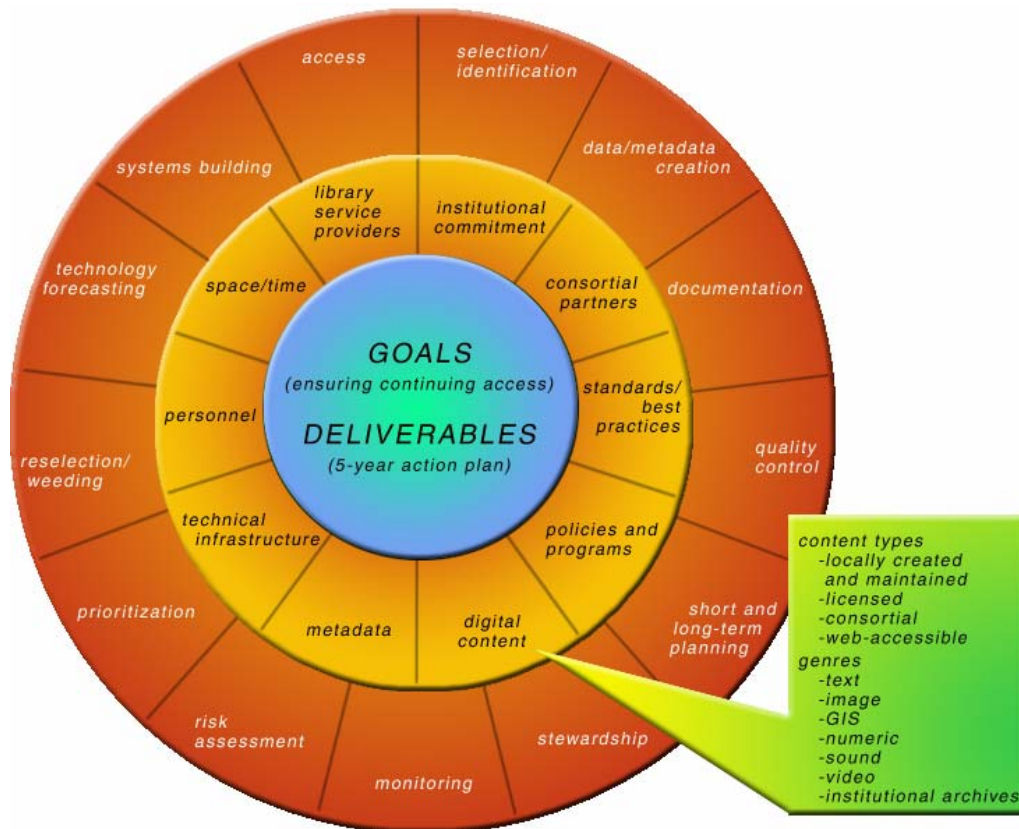
Figure 2: Digital Preservation Management: Connecting Goals (center) to Resources (middle ring) and Processes (outer ring)

Teaching Principles for the Workshop
1. Institutions need to understand the attributes of trusted digital repositories in order to build their own or to evaluate those built by others.
The recent Digital Library Federation survey reported that the greatest digital expenditure (40%) in member institutions in 2000 was for access to access commercial content, for which there are currently no reliable systems in place for long-term archiving. The survey also showed that subject bibliographers nearly always carry out the selection of these resources; many of them are not fully aware of the issues surrounding digital preservation. As commercial digital content grows, institutions must consider preservation needs in selection decisions, evaluate the emerging options for archiving content in trusted repositories, and assess the preservation claims of content providers. They will need to stay current with efforts elsewhere. For instance, the U.K. National Electronic Site Licensing Initiative for electronic journals (www.nesli.ac.uk) is in the process of mounting a revised license agreement that allows for collective archiving arrangements to be implemented by subscribing institutions.

Libraries and archives are also developing content or assuming direct responsibility for the long-term management of content entrusted to their care. Thus, one goal of the workshop will be to educate participants on current research and development efforts geared toward building formal digital repositories that will enable institutions to "ingest" digital content and preserve it for long periods of time.

In 1996, RLG and the Commission on Preservation and Access published a seminal document, *Preserving Digital Information, Report of the Task Force on Archiving of Digital Information* that recommended dialogue among "the appropriate organizations and individuals on the standards, criteria and mechanisms needed to certify repositories of digital information as archives." Since then work has moved forward through a number of large projects designed to develop a viable technical architecture for digital repositories. NEDLIB has focused on the deposit process whereby digital content can be ingested into a repository in a form that will ensure its long-term viability. CEDARS has focused on metadata requirements for effective preservation and on coordination of preservation activities among distributed repositories. The PANDORA archive has taken a pragmatic approach, collecting Web content in its original HTML format. The InterPARES project has focused on applying a specific set of archival principles to digital collections. In a project funded by the National Archives, the San Diego Supercomputer Center is investigating technology-independent methods for ingesting and managing digital collections. [11]

Cornell is directly involved in building trusted digital repositories for its own collections and for electronic journals owned by commercial publishers. Cornell's Mann Library has assumed de facto archival responsibility for USDA datasets and reports and maintains an online repository of GIS data and metadata for New York State, the Cornell University Geospatial Information Repository. Data downloads at CUGIR average over 5000 per month. In addition, Anne Kenney serves on the international working group sponsored by RLG and OCLC on Digital Archives Attributes, whose report will be released in September 2001. Principally drafted by Kelly Russell, the report defines the characteristics of reliable

---

[11] *Preserving Digital Information, Report of the Task Force on Archiving of Digital Information* (RLG and CPA, 1996); NEDLIB, www.kb.nl/coop/nedlib; CEDARS, www.leeds.ac.uk/cedars; PANDORA, pandora.nla.gov.au; InterPARES , interpares.org ; San Diego Super Computer, www.sdsc.edu/NARA.

archiving services for heterogeneous research collections and identifies tools that support institutions as they seek either to build their own archiving capacity or contract with third-party services.[12]

2. Institutions need to understand and implement guidelines, standards, and best practices in creating and managing digital assets.
In the past few years, a number of organizations, including Cornell, have sought to develop policy guidelines for creating rich digital master files that will retain their value well beyond the short, 3-5 year cycle of technology obsolescence in computing. For libraries in particular, the cost of retrospective conversion makes it imperative that digital files be created and managed using available standards and best practices. In particular, two organizations, PADI in Australia (www.nla.gov.au/padi) and the Arts and Humanities Data Service (ahds.ac.uk) in the U.K., and have made excellent progress toward organizing and disseminating information about best practices for preserving digital content.

As work progresses toward formal standards for digital repositories and preservation metadata, digital librarians and archivists need hands-on learning to develop policies and apply best practices to actual digital collections. Cornell's involvement in setting preservation standards through its work in RLG, DLF, and NISO (National Information Standards Organization), combined with its expertise in digital collection development, gives it a distinct advantage in bringing practical experience to cultural professionals.

By the time of the first workshop (June 2003), several important digital preservation projects will be completed, from which will emerge standards and best practices (see list of promising initiatives in Appendix A, tutorial description). Guidance will also come from the Library of Congress, which has been empowered by Congress through a $100 million special appropriation to develop a national program to preserve materials created only in digital form. Similarly, the National Archives will be securing sizeable additional appropriations to deal with preserving Federal electronic records. We will also monitor initiatives sponsored by the National Endowment for the Humanities, the National Historical Publications and Records Commission, the National Science Foundation, the Digital Library Federation, the Joint Information Systems Committee-Digital Preservation Program, the British Library, the Arts and Humanities Data Service, and the National Library of Australia.

National and international standards that relate to digital preservation, file formats, naming conventions, compression, storage, media, and text encoding, etc. will be monitored throughout the project via several print and electronic publications and listservs such as ANSI OnLine: Catalogs and Standards Information; Custom Standards Services; Internet Engineering Task Force; National Standards Systems Network; Digital Libraries Research Forum; Digital Librarianship; ImageLib; and PADI Forum.[13]

3. Institutions must appreciate organizational and technical risks associated with digital assets.
Cornell is an acknowledged leader in research on the technological risks facing digital files. Through its risk management investigation and related work, Cornell staff has gained substantial experience in analyzing the trade-offs that exist in choosing formats for access and long-term preservation. When institutions turn to proprietary formats such as PDF, digital librarians and archivists need tools for

---

[12] Geospatial Information Repository, cugir.mannlib.cornell.edu/; Digital Archives Attributes, www.rlg.org/longterm/attribswg.html

[13] ANSI Online (http://Web.ansi.org/public/std_info.html), Custom Standards Services (http://www.cssinfo.com/index.html), Internet Engineering Task Force http://www.ietf.cnri.reston.va.us/html.charters/wg-dir.html
 National Standards Systems Network (http://www.nssn.org/), Digital Libraries Research Forum (DigLib) (listserv@infoserv.nlc-bnc.ca) , Image Lib listserv@listserv.arizona.edu); PADI Forum, listproc@nla.gov.au; and Digital Librarianship (listserv@sunsite.berkeley.edu).

analyzing the impact of such formats on the quality and consistency of master files in creating new content and in migrating existing resources to them.

A key effort of Project PRISM is the Digital Information Longevity Study. Anecdotal evidence suggests cultural institutions are losing digital information due to technical and organizational threats, including obsolescence of various technical components, incomplete documentation, and lack of resources dedicated to preservation. But there are no systematic quantitative or qualitative studies that document the extent and rate of loss. The longevity study will analyze and document loss and evaluate its significance to the usability and value of digital information. An understanding of leading causes of loss will help to prioritize preservation strategies and will facilitate risk assessment and the development of preventive measures.

The role that redundancy plays in data protection is being addressed through Cornell's participation as a beta test site for the Stanford LOCKSS project (Lots Of Copies Keep Stuff Safe: lockss.stanford.edu/). LOCKSS software enables libraries to run Web caches for specific journals. These caches collect content as it is published and are never "flushed." By cooperating in a peer-to-peer network, libraries can detect and repair damaged or missing pages. The caches run on generic PC hardware using open-source software and require almost no skilled administration, reducing the cost of maintaining a journal in the state it was published. LOCKSS is currently being tested at 40+ libraries worldwide with the support of 30+ publishers.

4. Institutions need to mainstream digital preservation into their organizational mission.
From the early 1990s, many have argued that digital preservation requires attention to all phases of the information lifecycle, from content creation to management to access. As libraries devote more resources to digital collections, preservation must be considered a mainstream activity and not an isolated function with minimal impact on the organization. The DLF survey previously mentioned suggested that libraries are reallocating substantial resources from traditional operations, in addition to seeking new funding and grants for digital collections. Given the long-term costs of digital archiving, it is important that libraries plan for preservation needs as they invest in digital infrastructure.

Cornell is well positioned to address the planning needs in digital preservation, based on its historic effort to make digital creation, and now archiving, mainstream functions within the library. Experience gained through large projects such as the Making of America, cdl.library.cornell.edu/moa/, and, more recently, the National Science Digital Library, nsdl.cornell.edu/projects/NSDL-Cornell.htm, will no doubt be valuable to other librarians seeking to integrate digital projects more closely with library functions. Cornell will also apply the business risk model described in the report *Managing Cultural Assets from a Business Perspective,* by Laura Price and Abby Smith.[14]

5. Institutions must understand and develop responses to risks associated with digital assets over which they have no form of control.
Most digital preservation research focuses on requirements for establishing digital repositories and ingesting content or on managing the creation of digital content to ensure its long-term viability. But research institutions are increasingly concerned about resources they neither own nor control. University and scholarly communities have similarly come to depend on information resources managed at varying degrees of formality, from strictly controlled databases to Web resources managed in a mostly informal way. Can risks be identified and addressed at various levels of granularity and in a distributed, non-hierarchical system?

---

[14] Laura Price and Abby Smith, *Managing Cultural Assets from a Business Perspective,* Council on Library and Information Resources, March 2000.

Project Prism is investigating distributed models for digital preservation, in which information resources can be managed along a spectrum, from the highest possible level (a formal repository) to the lowest level (the unmanaged Web). One of our principle goals is to show how the integrity of informal collections, especially Web resources, can be raised at minimal cost, using automated routines for monitoring and validating files according to policies established by institutions wishing to preserve them. Our eventual goal is to create archiving tools that will enable libraries, archives, commercial database providers, scholarly organizations, and individuals to manage different sets of risks affecting the same resources, by coordinating policy enforcement across many collections and repositories. For example, an artist might insist on a very fine description of her work, specifying at a high level of detail the programs and even the hardware used to render a document. A library, by contrast, might choose to describe events at a much coarser level, to inexpensively monitor threats to whole collections rather than to individual objects. Archives, for their part, might regard any change in the contents of a database to constitute a threat event, whereas a commercial database provider may regard such changes as normal and desirable. Our basic framework accommodates all of these different values, while also providing a consistent basis from which users can measure the cost and outcome of preservation policies and actions. In a world of rapidly growing multimedia collections, we believe that all types of organizations will have to strike a fine balance between the desired level of integrity and the long-term costs of managing digital content. The tools Prism has been developing for automated policy enforcement, metadata harvesting, and routine file monitoring make it possible to envision a variety of services that could be used to manage information at varying levels of control. Workshop participants will use the knowledge generated by this project to analyze threats to information integrity and to plan for corrective action as a regular function of a digital preservation program.

Structure and Delivery of the Curriculum
Adopt Key Principles of Adult Education
The success of the proposed curriculum depends not only on the content but also on the underlying pedagogy and delivery method. There is tremendous value in developing a teaching approach that fits the characteristics and learning needs of the target audience. Our target audience includes adult professionals who have been charged with developing digital preservation programs in their home institutions. We will incorporate key principles and conditions that facilitate adult learning—especially in the realm of professional development. Our approach to teaching and learning will build on the following premises.

- *Adults are autonomous and self-directed learners who need to be involved in learning*. The workshop will include sessions that allow participants to work in small groups and assume responsibility for their own learning. The hands-on labs and breakout sessions will provide participants with opportunities to reinforce and apply the information presented during the lectures and via the tutorial.
- *Adults connect learning to their own work experiences and past knowledge base*. A concerted effort will be made to draw on the experiences of the participants themselves. For example, during the lectures, presenters will ask participants to share their knowledge and experiences related to the topics being discussed. Group work will provide opportunities to share success and failure stories, and to build a common understanding.
- *Adults are goal and relevancy oriented. Learning has to be directly applicable to their work to be of most value*. The workshop will be designed to move from self-assessment to the development of digital preservation policies and practices. One of the goals of the workshop is the creation of a comprehensive preservation plan appropriate to the distinct circumstances of the home institution.

Limit Each Workshop to Twenty-Four Participants
We have adopted a managerial approach for this workshop, so we intend to create an environment that encourages participants to reflect on the value of information received and to develop decision-making skills consonant with their own institutional context. Limiting each workshop to twenty-four participants

will encourage individual participation and constructive group dynamics. It will also enable Cornell to offer direct hands-on experience. Participants will work individually and in small groups throughout the week.

Because enrollment will be limited, formal selection criteria must be developed and communicated to potential attendees. The ideal candidate will be an individual in a position of responsibility to develop or influence an institutional digital preservation strategy. Applicants who can demonstrate institutional support (e.g., a letter of commitment by their director) will be given priority. Timely registration for the workshop will also be considered. Participation will be limited to two staff members from the same institution (additional staff members will be placed on the waitlist).

Utilize a Variety of Learning Methods
The workshop will incorporate five key teaching and learning formats, including presentations, breakout sessions, individual and small group work, labs and exercises, and group discussions. The goal behind using a variety of teaching methods is to accommodate individual learning differences. In addition, the workshop instruction team will consist of eleven instructors plus a keynote speaker, offering participants an array of teaching styles. As shown in Figure 2, Cornell's current workshop series has a heavy lecture component (45%). Our vision for the proposed workshop is an even distribution of the five key instruction methods.

▮ (red)	Presentations
▮ (green)	Case Study
▮ (purple)	Breakout Sessions
▯ (white)	Individual and Small Group Work
▮ (black)	Labs and Exercises
▨ (gray)	Group Discussions

Current Workshop   Instruction Methods
    Proposed Workshop
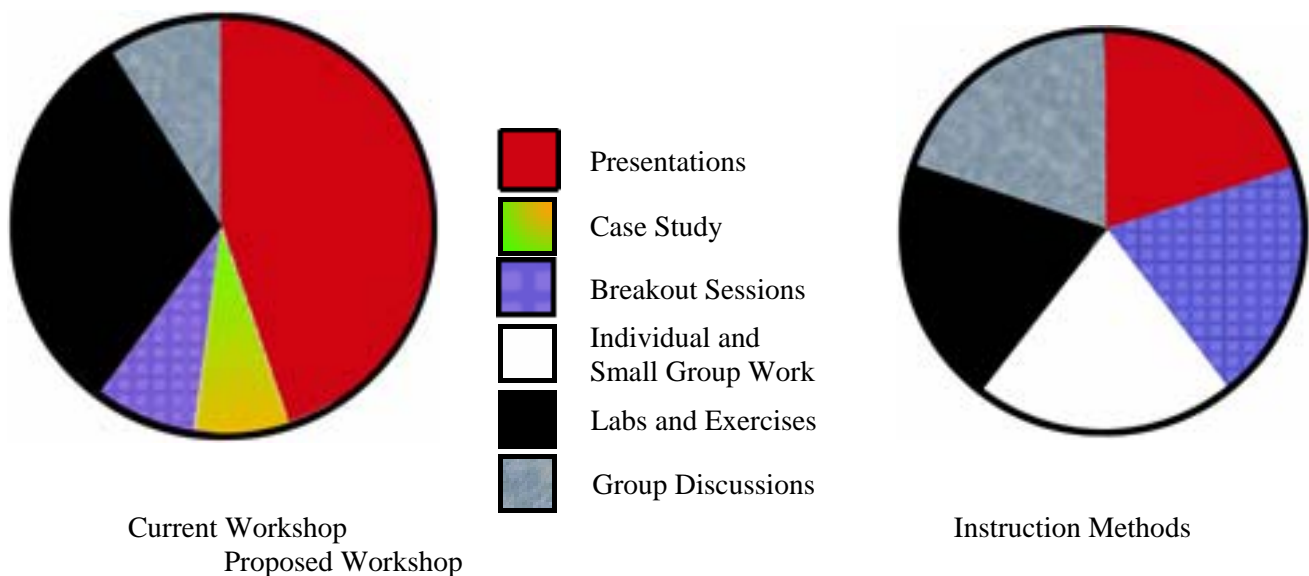
Figure 3: Comparing Instruction Methods: Current Workshop vs. Proposed Workshop

Take Advantage of Distributed Learning Opportunities
With the increasing popularity of the Web, distance education is turning into a beneficial and effective option for professional development. There are many advantages to the tutorial:

1. The tutorial will be accessible 24/7/365 in several languages, without fees or user restrictions, thus reaching a wide audience in need of elementary digital preservation information. A printable PDF version will be offered to those with limited connectivity.
2. Because the tutorial will not represent a formal course with predefined beginning and ending dates, users will have time and place flexibility to determine their own pace for learning. The tutorial will be organized into sections, allowing readers to repeat parts when necessary or to skip portions not of

interest to them. This component-based approach will also facilitate the staff's ability to maintain content currency.

3. The comments link will allow readers to send questions and suggestions to the teaching team via email. This will create a continuous stream of feedback for the instruction team as well as a direct connection for users.

4. The tutorial will serve as a pre-requisite for the workshop, ensuring that participants gain the same base-level knowledge prior to coming to Cornell. This will enable workshop participants to focus on issues and processes best addressed in a limited enrollment seminar. The focus on problem solving in the last workshop series was greatly enhanced by participants having completed the digital imaging tutorial prior to coming to Ithaca.

In addition to the project staff's expertise in creating Web-based tutorials, we will rely on the services offered by the Office of Distributed Learning, which provides administrative leadership for distributed learning in Cornell University Library. This office extends library services and programs into technology-mediated instruction and learning, conducts needs assessment to identify new service areas, and coordinates the development of innovative services and programs to support faculty, students, and administrators involved in distributed learning. The office explores the role of the library in managing and preserving digital course materials, and developing best practices and policies to support digital asset management.

Evaluation
Evaluation is critical to the success of this new education and training program and will take place on an ongoing basis. Various evaluation methods will be used to assess how well the curriculum objectives are achieved; to collect input for the continuous revision and refinement of the workshop and the tutorial; and to verify adherence to the principles identified in the
Methodology and Standards section.

We will evaluate the tutorial and the workshop at several stages. The first formal evaluation will occur prior to offering the first workshop, and will be conducted with the assistance of the project's Advisory Committee (see Project staffing). In addition, five outside experts will evaluate the tutorial in terms of content, usability, and presentation. Tutorial users will also be invited to complete an on-line questionnaire. Appendix G contains a sample evaluation form that was used with the digital imaging tutorial. This initial tutorial evaluation will be followed by periodic review of the server-generated statistics and users' comments throughout the course of the project. The evaluation results will be reflected in the quarterly updates and improvements to the tutorial.

During each workshop, instructors will respond to participants' feedback by implementing minor adjustments throughout the week. At workshop's end, participants will complete a formal written evaluation (see Appendix G for the form used in the current workshop series). The instructors will meet thereafter to review these evaluations in preparation for the next workshop.

The project director and coordinator will conduct a needs assessment at the end of the project. This assessment will include a systematic review of the evaluations from the previous workshops, an on-line survey sent to alumni, telephone interviews with a selected group of respondents, and a focus group discussion for the teaching team. This needs assessment will be shared with the Advisory Committee, which will assist project staff in determining whether to continue the digital training program and on what basis. The results of this assessment will be included in the final project report to NEH.

The success of this educational program will be measured in a number of ways. The first goal is to develop and present a base-level digital preservation tutorial and 8 weeklong intensive workshops. This goal will be met if all workshops are held, the tutorial is released, and users experience no major difficulty

in connecting to it. A second goal of this program is to ensure that these educational offerings meet the needs of their consumers. This goal will be met if workshop and tutorial evaluations average 3.5 or higher on a scale of 1-4, with 4 representing excellent. The third goal of this program is to encourage the development of effective short-term digital preservation programs. Project staff will consider this goal reached if 50% or more are able to make at least one significant change to protect digital assets at their institution within the course of one year following the workshop; if 30% are able to convince their institutions to adopt or significantly revise formal policies for digital preservation; and if 10% or more are able to implement such programs within two years of their participation in the workshop.

IV.   PLAN OF WORK

Cornell is seeking support to conduct a three-year program that will result in the development of an online, self-directed tutorial and the presentation of 8 intensive workshops to be held in Ithaca. The first workshop will serve as a beta test, and will be presented in June 2003 to the preservation officers and technical staff from the eleven New York State Comprehensive Research Libraries. The "Big 11" will seek tuition and travel support from the New York state Coordinated Preservation Program. One additional workshop will be held in 2003; three in 2004; and three in 2005.

1.   Tuition Support Justification
NEH support is also sought to provide partial tuition stipend, thus enabling a broader participation by representatives from a range of non-profit cultural research institutions. As described in the budget narrative, Cornell will charge a registration fee to all participants. This figure is competitive with registration fees charged for other workshops, especially when one considers the length of the workshop and the intensive hands-on approach taken. Without NEH support, each workshop participant would have to be charged to cover non-contributed costs. Clearly these costs could be borne only by a limited number of individuals from privileged or for-profit institutions. Additionally, the costs of preparing and maintaining the tutorial can be partially offset by the registration fee, benefiting many individuals beyond those attending the workshop.

2.   Project Implementation
The Plan of Work is summarized below, and Appendix E provides a Gantt chart of the outlined schedule. Staff activities and commitment are described in the *Staffing* section.

Refine and Prepare Curriculum Support Materials (Months 1-8)
- Refine the content, objectives, and learning outcomes of the workshop and tutorial based on changes occurring between grant submission and project implementation
- Identify additional staffing needs and skills required to update the curriculum (e.g., publications to read, conferences/meetings to attend, technical training)
- Assemble resources required for successful delivery of the curriculum: relevant print and electronic publications, Web-based examples, sample projects to be referenced during teaching.
- Incorporate a variety of teaching methods and learning strategies to diversify the delivery method
- Review/refine guest faculty outlines and participant exercises.
- Confirm lineup of keynote speakers and make necessary travel and other arrangements.
- Develop alpha version of tutorial
- Seek input on tutorial design and content from the Office of Distributed Learning, the Academic Technology Center, and the Human-Computer Interaction Group.
- Develop a responsibilities check list for the workshops (see Appendix F)

Identify Equipment Needs and Acquire/Purchase/Install Hardware/Software (Months 1-8)
- Reassess hardware, software, supplies, and networking needs to support the curriculum
- Acquire and install hardware and software to form the technical infrastructure

- Identify and secure older versions of operating system and applications software
- Negotiate for use of older software and shareware for emulation, Web harvesting, and simulation (e.g., Babbage Institute, software companies, Nordic Web Archive, Stanford)
- Acquire older versions of Macs and PCs
- Train staff in use of equipment and software
- Prepare curriculum materials, programs, tools and lab exercises utilizing the equipment.

Announce the Workshops (Month 6)
- Determine workshop dates for the entire series and confirm them with the instruction team
- Develop criteria for admission decisions
- Set up a Web site to disseminate information about the workshop (including information on the tutorial, admission requirements, application form, workshop dates and the associated deadlines, partial tuition stipend, etc.)
- Announce the workshops in several print and electronic publications
- Select participants three months in advance of each workshop

Conduct Initial Curriculum Evaluation (Month 7)
- Conduct formal evaluation of the alpha version of the tutorial
- Hold Advisory Committee meeting to review tutorial and workshop content, organization, structure, and teaching methods used
- Assess and evaluate the skills of the teaching and support team in light of these evaluations
- Refine tutorial and workshop based on these evaluations

Implement Education and Training Program (Months 9-35)
- Release beta version of tutorial to first three workshop participants in May 2003
- Offer eight workshops: June, August, and October 2003; May, July, September 2004; May and July 2005
- Refine workshop and tutorial based on continuing evaluation and participants' feedback
- Release tutorial for public use (December 2003)
- Release foreign language versions of tutorial (April 2004)

Conduct Ongoing Curriculum Evaluation (Months 8-35)
- Design evaluation questionnaire to be administered at the end of each workshop
- Design an online evaluation form to be completed by those using the tutorial
- Conduct on-going evaluation based on Web statistics and user comments.
- Analyze the comments/suggestions of workshop participants
- Arrange post-workshop instruction team meetings for self-evaluation of individual sessions and the full curriculum

Refine and Update the Workshop and Tutorial (Months 10-36)
- Provide continuing education opportunities for the instruction team for curriculum update
- Add new components to the workshop and tutorial to reflect changing digital technologies
- Revise the tutorial on quarterly basis
- Implement findings of the evaluation process to improve and refine the curriculum
- Respond to tutorial user inquiries in a timely fashion

Conduct Needs Assessment and Prepare Final Report and Recommendations (Months 35-36)
- Conduct on-line survey of workshop participants and selected users of the on-line tutorial
- Conduct telephone interviews with a selected group of respondents to find out how they translated the knowledge and experience gained during the workshop into practice. Inquire as to the on-going need for such a program

- Compare the workshop with other digital workshops to determine whether it continues to fill an educational need
- Consult granting institutions, such as NEH, NHPRC, and IMLS, about training needs for grant recipients via phone interviews with grants officers
- Hold final Advisory Committee meeting to discuss the findings of the needs assessment process and to form recommendations for future efforts
- Prepare and submit the final report to be submitted to NEH

V.   STAFFING

Anne R. Kenney, Associate Director of the Department of Preservation, has been involved in digital research for over a decade. She initiated the various digital training workshop series, and has written and spoken widely on the use of digital imaging for preservation and access. She will serve as the Project Director, responsible for fiscal management and administrative guidance. She will oversee the preparation and delivery of the curriculum, will serve as one of the principal instructors, and will undertake the needs assessment and prepare the final report.

Nancy McGovern, Coordinator, Digital Imaging and Preservation Research Unit, Department of Preservation, facilitates digital research, publications, and training projects. McGovern has over 15 years of experience in electronic records research and most recently authored the Cornell University Electronic Student Records Systems Project Report (2000). She will serve as program coordinator, managing curriculum development, including the tutorial and workshop products. As one of the two principal instructors, she will devote considerable time to preparing, teaching, and revising the workshop. She will oversee the quarterly updates of the tutorial, and participate in the workshop evaluation and final report.

Peter Hirtle is the Co-Director of the Cornell Institute for Digital Collections, and Associate Editor of *D-Lib Magazine*, a monthly magazine about innovation and research in digital libraries. Previously Hirtle worked for the National Archives first for the Technology Research Staff (where he contributed to its most recent study of digital imaging), and then as co-director of NARA's Electronic Access Project. He serves on the advisory boards of the NARA/San Diego Super Computer and InterPares projects. He will serve as principal advisor on curriculum content and will develop and teach sessions on legal considerations and licensing issues.

Richard Entlich is Digital Projects Librarian in the Digital Imaging and Preservation Research Unit. He does research, writing and instruction focusing on the technical aspects of digital imaging and preservation. Prior to coming to the Department, he worked at Cornell's Mann library on a variety of digital imaging and electronic publishing initiatives. Entlich will oversee the installation of equipment and software and the development of programs, exercises, and simulations used in the tutorial and workshop. He will develop portions of the tutorial and lead several sessions in the workshop, including ones on technical issues and preserving Web resources.

Peter Botticelli is Digital Projects Librarian in the Digital Imaging and Preservation Research Unit. His current focus is on digital preservation and the organization of digital libraries. He has specialized in electronic archives, including a two-year study of records appraisal in collaborative work environments. His background also includes research and writing in business history and the history of technology. He will develop portions of the tutorial and lead several sessions in the workshop, including ones on managerial issues and preserving Web resources.

David Ruddy is Electronic Publications Specialist, Cornell Institute for Digital Collections. His current focus is on implementing metadata standards for resource discovery and use, including the Encoded Archival Description (EAD), Text Encoding Initiative (TEI) and Extensible Stylesheet Language (XML).

He has been instrumental in developing text-based complements to the Making of America digital image collection. Ruddy will prepare a section of the tutorial and co-lead a session on preserving text and images.

Carla DeMello has been doing freelance work in graphic design and illustration for over fifteen years. DeMello oversees design and technical support for the Department's training and publishing efforts. She will participate in developing workshop exercises, creating the workshop Web site, and in designing/developing the tutorial. She will also be in charge of monitoring the tutorial and responding to questions and comments sent by the readers.

Barbara Berger, Assistant Director of Operations for the department, has spent the past thirteen years managing large-scale preservation microfilming and digital imaging projects. In recent years, she served as the chair of the ALA Preservation and Reformatting Section. She will participate in curriculum design and development, gather supporting material, monitor workshop arrangements, and serve as liaison with the translation service and outside speakers.

Mary Arsenault, Administrative Associate, Department of Preservation, will manage all workshop arrangements. Her responsibilities will include publicity, financial accounting, conference registration, arranging participants' accommodation, equipment purchases and rentals, printing and distributing course materials, and arranging conference rooms and catering. She will maintain and monitor project budgets.

A Programmer will work 100 hours during the first year, assisting in the development of the interactive tutorial (including setting up a server, writing CGI scripts for formulas, etc) and developing tools and simulations for the workshop. In years 2 and 3, the program will work 10 hours, updating workshop tools and the tutorial, maintaining the server, and to monitoring the CGI scripts used in the tutorial.

Guest Instructors
(See Appendix H for Résumés)

Nan Hyland, Project Manager of the Cornell University Geospatial Information Repository, will co-lead a session on GIS and numeric data.

Greg Lawrence, Government Information Librarian, Mann Library, will co-lead a session on GIS and numeric data.

William Kehoe, Programmer/Analyst currently developing the Cornell University Geospatial Information Repository and active in bringing Open Archives Initiative compliance to Cornell University Library's digital libraries, will co-lead a session on GIS and numeric data.

Oya Y. Rieger, Coordinator of Distributed Learning, Cornell University Library, will lead sessions on preservation metadata and digital asset management in the context of the virtual learning environment.

Susan Stinson, Curator, Belfer Audio Laboratory and Archive, Syracuse University Library, will lead a session on audio-visual recordings.

Key Note Speakers
(See Appendix I for Biographical Statements)

Neil Beagrie, Assistant Director in the Distributed National Electronic Resource (DNER).

Daniel Greenstein, Director, Digital Library Federation

Carl Fleischhauer, Technical Coordinator, National Digital Library Program, Library of Congress

Margaret Hedstrom, Associate Professor at the School of Information, University of Michigan

Clifford Lynch, Director of the Coalition for Networked Information (CNI)

Vicky Reich, Assistant Director, HighWire Press, Stanford University Libraries and Academic Resources, Director of the LOCKSS project

Jeff Rothenberg, Senior Computer Scientist, RAND Corporation.

Kelly Russell, CEDARS Project Manager, University of Leeds and principal author of the RLG/OCLC report on attributes of a certified archives

Colin Webb, Director, Preservation Services Branch, National Library Australia.
Advisory Committee
(See Appendix J for Biographical Statements)

Sarah Thomas, Carl A. Kroch University Librarian, Cornell University (Chair)

Caroline Arms, Technical Program Coordinator, National Digital Library Program, Library of Congress

William Arms, Professor of Computer Science, Cornell University

H. Thomas Hickerson, Associate University Librarian for Information Technologies and Special Collections, Cornell University Library

Carl Lagoze, Digital Library Scientist, Cornell Department of Computer Science.

Joy Paulson, Preservation Librarian, Mann Library, Cornell University **Dave Vernon**

Dave Vernon, Director of Special Projects, Office of Information Technology, Cornell University

VI.   DISSEMINATION

Upon notification of the award, we will announce the workshop series and the tutorial in several listservs[15] and library and archival journals.[16] We will also extensively advertise the program during several library, archival, and digital technology conferences such as the American Library Association Conference, the American Society for Information Science (ASIS) Conference, the Joint Conference on Digital Libraries, and the Society of American Archivists.

---

[15] The listservs will include: DIGITAL-PRESERVATION (UK), French Digital Information Preservation Group, Imagelib, Public-Access and Computer Systems Forum (PACS-L), Library and Information Technology Association List (LITA), Digital Libraries Research List, Archives Listserv, and the Preservation Administration Discussion Group Listserv.

[16] The print journals will include: *American Libraries, College and Research Libraries News, Microform & Imaging Review, SAA Newsletter, Archives and Museum Informatics.* The electronic journals will include: *Ariadne, D-Lib Magazine, RLG DigiNews, Digital Library News, Knowledge Management World, The Public-Access Computer Systems Review.*

We will maintain the Web-based tutorial on a 7-day, 24-hour schedule, with frequent maintenance checks to make sure that it is robust and functioning properly. We will also gather usage statistics to monitor how often and for how long the tutorial is being used. If necessary (slow response time), we will identify a mirror site to alleviate access time.

In the last two months, we will develop strategies for maintaining the program, perhaps partnering with other institutions. Cornell is committed to continuing the tutorial for at least several years beyond the life of the grant. We will consider offering additional workshops specifically aimed at instructors in other areas of the world. The goal would be to prepare them to deliver the course in their own countries. Resources for this "workshop in a box" would be translated into the appropriate languages for the targeted audience. This approach is currently being investigated for the digital imaging workshop series for use in Latin America and the Caribbean.