



The Preservation of Digital Content

Michele V. Cloonan and Shelby Sanett

abstract: The authors are conducting a three-part study to evaluate current trends in the preservation of digital content, with an emphasis on electronic records. The study emanated from the authors' work on the Preservation Task Force of the International Research on Permanent Authentic Records in Electronic Systems (InterPARES) project. This article incorporates the findings of both the survey and individual key-informant interviews that we conducted from August 2001 through February 2003, as round 2 of the study. Round 2 builds on the 2000–2001 round 1 survey that sought to identify and describe strategies for preserving electronic records. In this second round the authors found that progress has been made in some areas while it still lags in others.¹ The full study consists of three phases: round 1 identified and surveyed 13 institutions, projects, and programs in North America, Australia, and Europe. Round 2 surveyed eight of the 13 institutions again to follow up on their progress. Additionally, we interviewed 18 key informants, including archivists and librarians. In round 3 the authors will each conduct one case study drawn from the survey participants in rounds 1 and 2. By the end of the three rounds, the authors will have studied a continuum of activities (over a six-year period) that constitutes a range of digital preservation strategies. The study will have charted the change in technological developments over this period—developments that have occurred in our survey institutions to meet the requirements of their mandates to preserve digital content for as long as needed.

Prologue

Preserving authentic records in electronic systems is the Holy Grail for archivists. We are compelled to find solutions for problems that we cannot yet fully grasp. Not only is the Grail itself elusive, but the path to it is still largely untrodden. The first challenge is identifying what *is* a record, the second is appraising it, and the third is accessioning and preserving it. Other roadblocks include technological obsolescence, storage media fragility, the manipulability of electronic records, costs, and intellectual property rights issues. Then there is the Grail itself—the records that continue to change forms. In a short time we have moved from managing mostly flat files to man-

portal: Libraries and the Academy, Vol. 5, No. 2 (2005), pp. 213–237.

Copyright © 2005 by The Johns Hopkins University Press, Baltimore, MD 21218.



aging multi-media, dynamic, experiential, and interactive records. We do not know what documents will be like in the future or how will we access and use them.

We continue this quest, of course, because we have a responsibility and a mandate to preserve records for as long as they may be needed. Without a modern-day Galahad, we need to pursue a variety of strategies. The underlying question remains: How do we preserve digital content that is reliable, authentic, and accessible over time?² The purpose of this study is to identify approaches to and ideas about digital preservation. We discuss the findings of part 2 of our study to document extant practices as well as to solicit expert opinions on the digital preservation dilemma.

Definitions of Key Terms

Our research deals with phenomena from two distinct but overlapping fields: library practice and archival practice. Practitioners and scholars in these fields deal with similar or identical sources, but with slightly different perspectives. For example, academic librarians tend to think about preserving information permanently, while archivists usually preserve documents and records for as long as they are needed based upon retention schedules. Also, while archivists maintain that “part of a record without regard to its evidential value is not worth keeping at all,”³ special collections librarians would say that keeping part of a record—especially if that is all that is extant—is better than keeping no part of it at all. Archival and library vocabularies are not always congruent and require defining.⁴

- A **record** is a document made or received and set aside in the course of a practical activity. Thus a record is not merely private information. As one of our key informants put it, a record has a “fixed content” that can be “re-presented in the structure or form in which it was born.” The preservation of authentic electronic records requires that the recordkeeping processes must also be maintained. Librarians tend to use this term more freely to refer to documents and other materials—“materials” itself being a library rather than an archival term.
- An **electronic record** is a record that is created (made or received and set aside) in electronic form.
- An **authentic record** is a record that is what it purports to be and that is free from tampering or corruption.
- A **digital record** is one that now exists in electronic form, though it may or may not have been created in electronic form. A digital record may have been created on paper and digitized later. Subsequent digitization may change its “recordness.”

To these terms, we add:

- A **digital component** is a stored digital object that is necessary to reproduce an electronic record (or other digital asset).⁵
- **Digital assets and digital content** refer to all types of information, text, graphic, image, and multimedia (authors’ definition).



Background

This three-part study grew out of our participation in the International Research on Permanent Authentic Records in Electronic Systems (InterPARES 1) project, a three-year international and multidisciplinary collaborative research effort that examined the authenticity of records that need to be preserved.⁶ As members of the Preservation Task Force, we⁷ were asked to document current strategies in use for preserving authentic electronic records. Our study aimed to inform the primary research objective of the task force: to develop a model of the processes necessary to preserve electronic records.⁸ The primary activities of that model are: to determine the preservation requirements, select the technologies, bring the records into the system, maintain the records over time and across platforms, retrieve the records, and output the records for access. In consultation with the other task force members, we designed a survey instrument and sent it to 15 institutions or programs (of which 13 responded). The survey questions were based on the research questions that were the focus of the InterPARES 1 Preservation Task Force:

- What methods, procedures, and rules of long-term preservation are in use or being developed?
 - a. Which of these meet the conceptual requirements for authenticity?
 - b. Which methods of long-term preservation need to be developed?
 - c. Which of these methods are required (or subject to standards, regulations, and guidelines) in specific industry or institutional settings?
- To these questions, we posed additional ones having to do with definitions, preservation techniques (e.g. refreshing, emulation, migration, robotics), and costs.

Context of the Study

The only study that is similar to this one was carried out by Neal Beagrie between our round 1 and the completion of this round; it included four national libraries and four multinational initiatives.⁹ Since the Council on Library and Information Resources and the Library of Congress commissioned his report as part of the latter's planning for its National Digital Information Infrastructure and Preservation Program (commonly referred to by the acronym NDIIPP), Beagrie's emphasis is on national digital preservation initiatives primarily in libraries rather than in archives. The difference between his approach and ours is that he aimed to produce "a high-level survey" that was not detailed in terms of practice or procedures.¹⁰ Our study, on the other hand, aimed to learn about practices. Since Beagrie does not include a copy of his survey instrument, specific comparisons are not easy to make. In our conclusion, however, we will compare some of his findings to ours.

Some progress has been made since 2000; institutions are beginning to think about cost issues and models, but they have been slow to develop digital preservation policies and plans.

The findings of round 1 of our study were reported in an article in *The American Archivist*.¹¹ In that study we found that a number of preservation techniques were in use



but that none of them could be considered meeting archival requirements for authenticity. It also revealed that, although they are developing technological processes to preserve authentic electronic records, almost every institution had deferred costing digital preservation processes and implementing digital preservation policies. Some progress has been made since 2000; institutions are beginning to think about cost issues and models, but they have been slow to develop digital preservation policies and plans. As one respondent from round 2 observed about his institution, "As long as there is no plan, the risk will be that preservation will be ad hoc, inconsistent, and not embedded in the organization."¹²

In round 2, we refined and/or expanded our questions based on our findings in round 1 (appendix 1, http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v005/supp_open/5.2cloonan_appendix01.html). Since the results in round 1 revealed that the field is still fluid, we decided to continue to monitor the programs we surveyed in round 1 and to place their activities into a broader preservation context (appendix 3, http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v005/supp_open/5.2cloonan_appendix03.html). For the survey portion of this research, there were only eight survey participants and 18 interviewees. This is too small a sample to allow us to generalize; but we are able to identify trends or potential trends, because they nevertheless yielded a good deal of excellent qualitative information. These participants represented Europe, North America, and Australia. Lack of resources caused us to limit the survey and interviews. Though we could not include Asia, the Middle East, and South America, we are aware of digital initiatives in some of these areas. We also limited interviews to participants who spoke English.

One technique for gathering information is to interview experts in a field to benefit from their many years of experience and gain a perspective that is broader than that of most preservation managers. Interviewing "key informants"¹³ gave us confidence that the results of the surveys could be placed within the context of the preservation field overall, because we would be able to compare the survey results to our key informants' more broadly informed perspectives. We interviewed 18 key informants with diverse backgrounds who worked in a variety of settings internationally such as universities, government archives, foundations, and professional organizations (appendix 4, http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v005/supp_open/5.2cloonan_appendix04.html). We selected key informants because they presented several noteworthy characteristics:

- Seventeen of the 18 have a minimum of 15 years of experience in their profession and have experienced the transition from analog to analog plus digital preservation; the remaining informant has over 10 years experience.
- We selected people with wide-ranging perspectives:
 - They are faculty and/or researchers.
 - They are consultants/researchers.
 - They are funders from the private or public sector.
 - They work in government agencies.
 - They have preservation experience either directly in the field or as educators, funders, or collaborators of practitioners of the field.



- They represent other organizations that are consortial or that sponsor research.
- We selected some of the informants, because we thought they might provide thoughtful answers about the field based on their experience.

All told, between survey respondents and key informants, we spoke with 26 people from Australia, Canada, the United States, and Europe. Still, since archival and library models differ from country to country, our view is necessarily United States-centric.

Methodology and Its Limits

Our study uses a mixed-methods approach: survey and follow-up interviews (round 1); survey, follow-up interviews, and key-informant interviews (round 2); and case studies (round 3). Shelby Sanett alone, Michele Cloonan alone, or Sanett and Cloonan together conducted the survey and key-informant interviews. The surveys and follow-up interviews to the surveys took place between September 2001 and February 2003. The key-informant interviews took place between August 2001 and February 2003. Follow-up conversations with key informants took place throughout 2004.

Many researchers in the social sciences use mixed methods techniques when it is appropriate to link quantitative and qualitative methods. John Creswell describes mixed methods this way:

[The] approach is one in which the researcher tends to base knowledge claims on pragmatic grounds (e.g., consequence-oriented, problem-centered, and pluralistic). It employs strategies of inquiry that involve collecting data either simultaneously or sequentially to best understand research problems. The data collection also involves gathering both numeric information (e.g., on instruments) as well as text information (e.g., on interviews) so that the final database represents both quantitative and qualitative information.¹⁴

The mixed-methods approach best suited our overall strategy to evaluate current trends in and perspectives on the preservation of digital assets. We began the research by gathering data about digital programs and practices (round 1) then repeated the survey with a subset of the first survey participants (round 2). At the same time as we administered the survey in round 2, we developed a detailed view of digital preservation issues through our interviews with key informants. In the process, we collected closed-ended quantitative data and open-ended qualitative data.

Those who use qualitative methods have examined the relationship between researcher and respondent. Grant McCracken writes, "This intimate acquaintance with one's own culture can create as much blindness as insight. It can prevent the observer from seeing cultural assumptions and practices."¹⁵ In this study, familiarity with the field led the researchers to some of the interviewees. On the other hand, this same familiarity may have prevented the researchers from asking more questions when they should have.

Survey and Follow-Up Interviews

We designed the survey questionnaire to track specific institutional practices, especially in three overlapping areas: staffing, cost, and policy—though even in these areas the



emphases were intentionally different.¹⁶ Further, in passing, the interviewees spoke briefly about intellectual property rights and how institutions are dealing with them. The survey was pre-tested by members of the InterPARES Preservation Task Force. Then we administered the surveys while key-informant interviews took place. At the outset, we gave equal weight to the two strategies; but, in the final analysis, the study was more qualitative because of the small survey size, and because we conducted follow-up interviews with the survey participants.

We sent the surveys to the appropriate contact at each institution, and each person returned the completed survey either electronically or via post. We noted where we needed clarification or amplification, and we interviewed each survey respondent by telephone or in person. We taped the interviews, and graduate students from the School of Information Science and Policy at the State University of New York (SUNY) at Albany transcribed the tapes.¹⁷ We made preliminary corrections and sent the transcriptions to the respondents for further corrections.

Interviews with Key Informants

We did not pre-test the key-informant interview instrument since we designed it to complement and amplify the survey questionnaire and responses. The key informants received the questions in advance of their interviews (appendix 2, http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v005/supp_open/5.2cloonan_appendix02.html); but we did not ask them to prepare answers, because the questions were designed to be points of departure that would generate discussion. Of the 18 interviews, 14 were face to face, and four interviews were on the telephone.¹⁸ All of the interviews were taped except one, during which the tape recorder malfunctioned. We interviewed the key informants individually as well as in groups of two or three. We selected participants for the groups who had a variety of working relationships. We tried to select group members whose experience would be complementary. In one interview, the informants worked for the same organization but in different parts of the country; another group was working on a joint research project; another served on a task force. Still another group's participants worked together in professional associations. As with the survey interviews, the tapes were transcribed by graduate students and then corrected by the interviewers and informants. Quotations from key-informant interviews appear here with their permission.

Notable

We obtained far more information than we could use in this study, so we concentrated on the themes that related most directly to the questions. During the key-informant interviews, the interviewees cited books and articles. Some of the readings they had found influential; others were just mentioned as relevant items that they had read recently. Since we had not specifically asked them about their reading, these spontaneous citations are of some interest. Who reads whom? Who cites whom? The answers to these questions appear in appendix 5, http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v005/supp_open/5.2cloonan_appendix05.html, in which we list all of the writings that were cited, who cited them, and the frequency with which each item was cited.



By using different methods and sources in this study—surveys, key-informant interviews, and a literature review—we sought to corroborate our findings for the themes we examined. By triangulating our data, we sought to cross-validate our small sample of institutions and individuals.¹⁹

Findings: The Survey

The following discussion follows the order of the questions in the survey questionnaire (appendix 1, http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v005/supp_open/5.2cloonan_appendix01.html). The respondents (appendix 3, http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v005/supp_open/5.2cloonan_appendix03.html) focused on their institutional collections rather than on their internal administrative or financial records.

Program and Policy

The responses to these questions focused mainly on programmatic issues rather than policy. We returned to questions about policy in the last section of the survey. We wanted to learn whether institutions had progressed in determining criteria to choose a particular preservation technique or, if they provided updated information, whether that decision was correct for their needs.

Since the first interview in 1999, seven respondents have reached the point of testing or evaluating preservation methods or techniques. All but one plan to disseminate results via Web sites, conferences, and publications. The remaining institution considers the information to be proprietary.

Specifics of Preservation Technique

The preservation techniques or methods that the institutions are using or exploring include emulation (3), e-Permanence/Normalization (1), Persistent Object Preservation (1), and migration (3). One project is exploring both migration and emulation.²⁰

Institutions chose preservation techniques in a variety of ways: internal research and development, internal proprietary software system development, “software systems . . . based upon data handling systems, local collection management, digital libraries, Web interfaces, and archival storage systems.”²¹ Other respondents indicated that the choice had not yet been made: “We are testing, as opposed to using, migration and emulation in the project. We are attempting to gain a deeper understanding both of cost issues and also of appropriateness of each technique for different categories of digital objects.”²² “In the test bed project—migration, emulation (not yet), and standardization (XML). No strategy has been chosen yet; they are all under exploration in the tested project.”²³ “We’re pursuing the development of persistent archives as the primary technique.”²⁴ One respondent is proceeding by “reference to best practice documents (e.g., the Cornell University report on file formats).”²⁵ Evaluative data on the efficacy of the preservation method/model are pending for most of the institutions surveyed.

Six respondents each indicated that a factor in choosing the appropriate preservation technique for their institution included its potential effect upon the intellectual integrity (e.g., authenticity and reliability) of the digital material. Steps taken or planned



in order to prove that the intellectual integrity of the digital material had not been compromised through the preservation process include “definition and documentation of the record’s ‘essence’; use of hashing techniques on XML source; [and] repeatability of the ‘normalization’ processes through the preservation of original bit streams.”²⁶ “Proof integrity (authenticity) [that] has not been compromised should be given by metadata about the preservation process and the management process in general. This should also be the case if in some way integrity has been compromised during the preservation

Responses indicate that institutions are taking steps to protect the intellectual integrity of the digital content, and they give the reasons for choosing appropriate preservation techniques. This indicates the rapid progress made in this arena.

process.”²⁷ “In terms of the test sites, we prepared information packages with a view to protecting them from undocumented changes by restricting access to both read and write operations.”²⁸ “Original bit stream preserved; all migrations will be checksum and regression validated. All migrations will be documented and new versions viewed as secondary manifestations.”²⁹

they give the reasons for choosing appropriate preservation techniques. This indicates the rapid progress made in this arena.

Responses indicate that institutions are taking steps to protect the intellectual integrity of the digital content, and

Cooperation

Insufficient resources compelled many institutions to develop strategic internal and external alliances or collaborative relationships with public and private sectors. Primary areas of collaboration included research, development, and testing.

Four institutions cooperated with other organizations to develop digital preservation initiatives—collaborating with archives, libraries, public companies, and other agencies. Collaborative efforts took place on local, national, and international levels. The collaborative efforts evolved partially out of multiple funding resources; one respondent reported, “The project was funded by eight different agencies. Each project involved collaborations with other institutions.”³⁰ Two respondents discussed various strengths and weaknesses attributable to the collaborative structure of the project. Strengths include complementary knowledge and/or experiences, multidisciplinary perspectives, breadth of efforts in data collection in digital libraries, data grids, virtual data grids, and development of the concept of persistent digital archives. There is tremendous energy in use of technologies from these research areas. Weaknesses cited include the multiple sets of competing standards for annotating information and knowledge as well as dependence on the knowledge of others. However, the respondent noted that that might be an ongoing issue, given the rapid developments in this area.³¹

The advantages of cooperation outweighed the weaknesses. The need to be cost-effective drives the development of strategic alliances; collaborative relationships among institutions are in themselves an effective strategy.



Staffing

We explored staffing issues, because we wanted to learn about the roles that formal training and on-the-job training play in digital preservation. Are formal curricula sufficiently able to identify and prepare students to acquire skill sets required for this type of work? Do employers rely on graduates of library, archival, and information science programs to staff their projects?

Regarding the requisite skills required to work on the preservation of electronic records, several institutions answered both yes and no with comments. Required skills that were identified include:

- Document analysis, XML, and imperative programming skills; micro-appraisal skills; document modeling and definition³²
- Record-keeping expertise, record-information analysis, IT knowledge, project management, and communication skills, among others³³
- Database, information management (XML, MIX), knowledge management (UML, DAML-S, model-based mediation)³⁴
- “At a minimum you need [professionals in] three [areas]. One is an archivist who knows what the requirements of the work are and the nature of the objects they’re dealing with. One is a computer specialist who can mount and maintain applications. And the other is an . . . ‘archival engineer’ who has a combination of knowledge from the archives side and from the IT side.”³⁵
- (1) Functional analysis: knowledge of and an ability to perform functional analysis of electronic records systems; (2) data management: knowledge of and ability to work with data management concepts and systems; (3) evolution of automation: knowledge of and ability to recognize the various stages of the evolution of automation (hardware, computing environments, among others); (4) software: knowledge of and ability to work with various types of software; (5) records management and document management systems: knowledge of and ability to work with records management and document management systems; (6) National Archives practices and approaches to electronic records: knowledge of and ability to apply the institutions’ practices and approaches; (7) viewer and file editing software: knowledge of and ability to use; (8) IM/IT concepts: knowledge of and ability to apply when working with e-records; (9) trends and best practices in ER archiving: knowledge of trends and best practices in other archives and in the IM/IT industry.³⁶

Four institutions noted that their organizations *could* find staff that had the required skills, although one commented, “not in sufficient quantity for the future”; and another commented that they “would need major staff development to locate and train staff.” Four respondents indicated that their preservation projects would draw staff away from other departments or projects.

We asked the respondents to describe the academic/professional background and job title/duties of their *full-time* preservation staff:

- Professional archivists: minimum is BA (honors) in history and graduate studies in IT³⁷



- The internal persons involved are archivists and people with IT background; the contracted people—IT, computer science, archival theory.³⁸
- Research interest in development of data, information, and knowledge management systems; expertise in databases, information repositories, knowledge repositories, XML, UML, RDF, Topic Maps, ontology development³⁹
- IT/computing/engineering/archives postgraduate level⁴⁰
- Professional archivists with IT training; master's degrees and one PhD; will be looking for an academic computer scientist, probably one holding a PhD⁴¹
- Training in history—others are geography, statistics, engineering, and computer science⁴²

Responses to the same question regarding *part-time* staff include:

- Information science graduate diploma and an interest in preservation and long-term accessibility to electronic records⁴³
- Expertise in C, C++, logic programming, database applications⁴⁴
- Administrative/secretarial⁴⁵

The majority of respondents reported that staffs receive training that supplements their project-based, on-the-job training. Types of training for staff include various continuing education courses offered by commercial providers, on-the-job training, participation in standards groups, review committees, conferences, and vendor collaborations. The frequency that staff updated their skills ranged from continually to more structured, quarterly training courses.

The respondents plan to contract out some of the major and/or minor components of their projects. Contracted items include software development, hardware supply and installation, development of training programs, programming, and various minor non-core aspects of the project work.

Responses from round 2 were much more detailed with respect to skills expectations than they were from round 1. This is consistent with evolving needs of those projects that have progressed since round 1. The overall impression from the responses indicates that library schools, archival programs, and computer science programs should be developing strategic alliances to meet those needs.

Technical Questions

To questions regarding the current status of their preservation practices, respondents offered a variety of perspectives. They carry out preservation activities in-house, including:

- Setting up a procedure for transfer from a government organization to the archive
- Creating an accessioning template
- Establishing digital object formats and markup languages
- Identifying the processing steps needed to convert to platform-independent representation
- Establishing a digital archive interface for loading records

They relayed information indicating the strengths and weaknesses of their institutions' preservation methods or techniques. They perceived their programmatic strengths as:



- Maintaining the integrity of preservation masters; use of public domain/open source technologies; and almost certain lack of reliance on outside businesses over time⁴⁶
- Defining a process that supports the migration of collections onto new technology, in which the standards used to characterize the data also change⁴⁷
- Developing archival requirements rather than addressing technical problems⁴⁸

Respondents perceived their programmatic weaknesses as:

- Still at early stages of development⁴⁹
- Necessarily dependent on commercial standards' becoming widely used for knowledge management, because the technologies will continue to change rapidly⁵⁰
- Dependent on the marketplace—"Most of what we are looking at doing is not currently available in the marketplace. The predictions are it will be there within a couple of years; that's one weak point. The other . . . is that to develop the system, we are going to need a lot more money."⁵¹

To the question "How are you storing the electronic records that have been preserved?" respondents mentioned the use of several procedures (or lack thereof):

- On online media with optical disc and magnetic tape back-up⁵²
- Not yet happening⁵³
- "We build a logical collection to store attribute values, aggregate digital objects into containers, and manage the storage of containers in archives through a data-handling system."⁵⁴
- Large-scale data servers⁵⁵
- "It hasn't changed yet, but it will change. What we've looked for is reliable storage media. Most records being preserved aren't used most of the time so you want to write them to something that will last awhile [be]cause it costs you a lot of money to copy it to new media. It's beginning to look like copying to new media can actually save you money, thanks to Moore's Law."⁵⁶ "So what we're looking at is a new equation which says, 'It's not so much that I want it to last for 10 years versus 5 years, I want it to be reliable for however long I'm going to keep it and that length of time factors in[to] how much it cost me to keep it on that medium versus to migrate it to a newer medium which is denser and faster and less expensive.'"⁵⁷

The needs for widely accepted standards for knowledge management and quality verification as well as a lack of sufficient financial resources were points made in both survey rounds.

Between rounds 1 and 2, new technological resources have become available for data storage, e.g., spinning hard disks. Programmatic strengths and weaknesses discussed in round 2 naturally illustrate a greater level of detail than provided in round 1. The needs for widely accepted standards for knowledge management and quality verification as well as a lack of sufficient financial resources were points made in both sur-



vey rounds. Implementing satisfactory quality-control procedures remains a goal for many institutions. This lack may be attributable to the constant evolution of hardware and software in the field.

Preservation-Related Costs

Increasingly, participants at conferences focus on the need of institutions responsible for preserving records to determine potential costs to preserve electronic records over time.⁵⁸ Issues include methods to determine costs, selecting appropriate cost elements, the use of cost-benefit versus risk-benefit analysis, development and analysis of costing frameworks and cost models, and the integration of this information with institutional decision-making strategies. Discussions lead to the need for policy development in this area largely because of the need for institutional support over time. Based on responses in round 1 of our survey and the increasing interest in this topic, we broadened the scope of our questions about cost activities. We also deepened them, hoping to learn what progress institutions had made in developing cost figures and policy related to the preservation of electronic records. We grouped questions into three broad categories: value, budget, and insurance.

Five of eight institutions (63 percent) surveyed have a general preservation policy that includes records in electronic form. In round 1, only three of 13 had such a policy (23 percent). Because the number of institutions surveyed in round 2 is smaller than those surveyed in round 1, this information should be considered only as a potential indicator of an emerging trend.

1. Value

Only one respondent could estimate the monetary value of the records in the archive. One respondent commented that, in a government archives, the records have evidential, business, and cultural values. We asked the respondents who were unable to assign a value to the records whether there was a problem measuring the value of the records in the archive. The discussion ranged from the theoretical to the practical, but there was no consensus on how to value records or whether they should be valued in a monetary sense.

When asked what the estimated value of the archive was, one institution responded: "The cost of acquiring the data."⁵⁹ Asked how that value was calculated, two respondents commented:

- "The value is the effect it has on whatever you're trying to do with the data in the future. The value of the data is the purpose you're going to use it for in the future."⁶⁰
- "There are cases where one could absolutely assign a specific value because the existence of a given record which the government will certify the authenticity [of] can make a court case—make or break it. But that depends on the use to which you're putting it."⁶¹

In both of these responses, the value of the records in the archives was closely connected to the function of the archives.



2. Budget

We asked whether respondents have an annual budget for digital preservation activities. Five respondents do; one does not. One respondent indicated that there is not one identified budget but rather parts of various budgets are used. Annual budget amounts ranged from USD \$300,000 to USD \$3 million.

When asked, “If you do not have a regular budget for digital preservation, or if the budget is demonstrably insufficient, does this create a risk that you can measure?” Two institutions replied in the negative. One commented, “No, but there is a huge risk, of course, for losing electronic records.”⁶²

As one might expect, three institutions indicated that problems could be resolved with additional funds. Another respondent pointed out, “There are always problems that can be resolved with more money, but equally there are always problems that will exist irrespective of how many \$\$\$ you throw at them.”⁶³

We asked what the impact would be if these problems persist. Responses were similar:

- “The impact would be that the Archives cannot properly preserve electronic records.”⁶⁴
- “It’s hard to quantify the value in future. The cost of the labor will still be there. The major challenge is finding funds to apply the technology to new disciplines.”⁶⁵
- “Loss of accessibility to the records.”⁶⁶

Asked whether the annual budget was sufficient for routine digital preservation activities—two institutions responded yes; two responded no. One commented that their institution did not know yet. Four institutions responded that their current budget could fund a project to preserve electronic records. Amounts allocated for this purpose ranged from USD \$100,000 to support the preservation of an electronic records repository to USD \$250,000 for research and development. These institutions considered the amounts to be sufficient for their activities.

One institution indicated how they calculated their cost estimates: “We have the luxury of additional projects that are developing the underlying technology (data handling system, information catalog management system, knowledge characterization). We can build a collection and archive it for the labor cost to support automated ingestion of the collection.”⁶⁷

3. Insurance

When asked whether the institutional holdings were insured, four institutions responded that they were not, and one said they were. Two government archives commented that their governments carried their own insurance so the collections are self-insured. None of the respondents that have uninsured holdings plans to obtain insurance.

When asked how the value of uninsured holdings was determined, none of the institutions responded.

Many cost issues appear in both rounds of the survey. As projects and programs mature, substantive data on actual costs to preserve electronic records should emerge as should new cost models.



Records Organization and Provenance

Because this section of the questionnaire dealt exclusively with archival issues, we are not discussing it here. To read the questions, see section I of the questionnaire in appendix 1, http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v005/supp_open/5.2cloonan_appendix01.html.

Preservation Plans, Actions, Procedures, and Policies

Under this rubric, we asked several questions. This area evolved from round 1 of the questionnaire in which it was evident that policy development was lagging behind technological advances. We wanted to learn whether there was less of a gap three years later—from the beginning of round 1 to round 2.

- Does your institution have a digital preservation plan?
Of the respondents, three did and three did not have such a plan; two did not answer the question. One participant's plan was undergoing revision at the time of the survey.
- If not, does *not* having a digital preservation plan create a risk you can measure? Will this risk have an impact on the institution or archive? A noteworthy response was, "A plan is in preparation at _____. As long as there is no plan the risk will be that preservation will be ad hoc, inconsistent and not embedded in the organization."⁶⁸

Of the institutions with digital preservation plans, the plans are readily available to staff, organizational management, and archive stakeholders. Respondents reported that staff, including management, review plans regularly.

- Can the institution's administration use the plan to understand how strategies to preserve electronic records fit into the operations of the archive or institution? Five participants responded: four answering positively, one negatively.
- Do you have a general preservation policy that includes records in electronic form?
Of the 6 responses, 5 participants answered affirmatively.
- If not, do you have a policy for reformatting, refreshing, migrating, and emulating data to newer technological platforms?
Only one institution has a policy.
- Please describe any policies you might have that relate to preservation of electronic records.
One participant was quite specific as to the scope of the policy: "The major policies are the specification of: standards to use for markup languages; migration policies (how frequently to track technology updates); access standards (what mechanisms to use for discovery of data); and conformance to emerging persistent archive standards."⁶⁹ A second respondent said that the policy was included in annual business plan targets.⁷⁰



Intellectual Copyright and Preservation

The present article focuses only on the issues raised for round 2. In our questionnaire we specifically noted that reproduction and copyright were to be deferred to round 3 (see appendix 1, M, http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v005/supp_open/5.2cloonan_appendix01.html).

However, as mentioned above, the area of ownership of intellectual property is such an important issue that some of the interviewees raised it in round 2. Since the topic arose in the interviews, a few words about it are appropriate here.

As Clifford Lynch has pointed out repeatedly, the most intransigent problem related to the scanning and subsequent preservation of information is about clearing property rights. He says,

It's not . . . fundamentally an economic problem of paying for scanning, but one about clearing rights to digitize material. Things remain in copyright for an astoundingly long time under the current legal regime. The libraries would have to clear rights book by book for every book published after, say, 1920-something. It's an incomprehensibly monumental and costly task—and a stupid one for them to have to undertake.⁷¹

Similarly, Michael Lesk has pointed out,

The problem is not that libraries or users are not willing to pay for intellectual property as much as it is a problem of finding an administratively workable and fair way of doing it. . . . Issues related to intellectual property law are the most serious problems facing digital libraries.⁷²

While this is a pressing issue for librarians, because most books and journals they acquire are still under copyright, archives also contain some collections with copyright restrictions. Since most of the key informants for this study were archivists, rights tended to be mentioned as one of *many* obstacles to preservation. Had we interviewed only librarians, it might have been more central to their discussions.

As noted in figure 1, several informants listed this as a critical issue, one calling it “the muddy waters of rights management,” an indication that the topic is endemic to *all* people working with information. A sustained discussion of this topic constitutes part of our work for round 3.

Findings: The Key-Informant Interviews

The key-informant interviews (appendix 4, http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v005/supp_open/5.2cloonan_appendix04.html) provided many perspectives on current preservation programs, activities, and research (see figure 1). They also represented broad, expansive, and even utopian perspectives, as is illustrated in the responses quoted below from the interview transcripts.⁷³

How do you define preservation and has your definition changed over the years?

- We (Cloonan and Sanet) reported in round 1 a shift in the definitions away from the artifactual approach that predominated for paper-based records to the view for digital records today that the aim is to preserve the ability to reproduce a



Figure 1
Key Challenges That Affect Digital Preservation
As Identified by the 18 Key Informants in the Interviews

Technological Issues

- Obsolescence of hardware
- Computer viruses
- Computer crimes
- Media and software deterioration
- The debate over preservation strategies, e.g., migration, emulation
- Disappearing information (both file formats and entire infrastructures)
- Interrelation problems (where are the boundaries of the work?)
- Maintaining relationships or links between digital content and digital metadata
- Metadata (definition; the need for crosswalks)
- Intellectual property issues (“the muddy waters of rights management”)
- Short length of life of media coupled with our lack of recognition that it is more than the media itself that needs to be preserved
- How to preserve through time the key characteristics of records

Strategic Issues

- Identifying what needs to be preserved
- Custodial problems (who is responsible for saving it?)
- Dependence on global context
- Economic issues (where does the responsibility lie for support?)
- All technology depends on technical support, which requires stable funding
- Archivists/librarians must keep making the case for preservation
- Lack of preservation awareness on the part of the public
- No clearly defined methods or best practices
- Lack of standards
- Lack of a research agenda
- Ignorance; e.g. educating administrators/users to understand that digitization is not preservation
- Management of the process
- Mindset (educating professionals to be pro-active)

Archival Issues

- Selection and appraisal (what to keep?)
- Deciding what is of value
- Conceptual issues (what is an archive? what is a digital archive?)
- Evidence and accountability
- Authenticity and reliability (trustworthiness)



record.⁷⁴ From the key-informant interviews, a broader perspective emerged. Two of the informants emphasized that preservation is stewardship. “I keep adding pieces to my original understanding of preservation and I don’t think of it in terms of ‘500 years’ or ‘techniques’ but rather more in terms of stewardship: what body of materials are we as libraries and archives responsible for from society’s perspective? . . . So I do not have any changing definitions of preservation. But I think that the stewardship rule does not change.” Another similarly said, “I see preservation as a function of stewardship and custodianship.”

- An archival scholar and educator said, “Speaking of ‘recordness’ and recreating ‘recordness’ in a digital environment: recordness relates to the qualities that make something a record, like having a fixed content and always being re-presented in the structure or form in which it was born, along with being able to recreate the context in which it was created. It also entails being able to demonstrate and track what has happened to the record over time; this includes what record-keeping management processes have been applied to it—being able to have an audit trail of that. . . . You can talk of preserving content, preserving the structure, preserving the context, and continually linking the records themselves to the related records of the record-keeping processes. There may be many implementation solutions in terms of the way that digital preservation is carried out.”
- We heard a more idealistic definition from one informant: everyone must be a preserver in the digital world. “I would like to make a pitch for preservation at a local level, at the individual level. I’d like each individual to be able to preserve his or her digital materials and contribute to some kind of archive, a preservation institution.”

Under “Staffing, Education, and Training,” we wanted to know whether the informants felt that LIS programs are up to the task of preparing students to work with (and preserve) digital records.

- In a wide range of responses, two asserted that the most important skill to have is critical thinking; they both felt that a good liberal arts education provided that grounding. One of them also emphasized the importance of archival training: “When we started to do the work on digital image quality, my archival experience was much more important to me than any technical knowledge. It was that profound love of the document that kept me engaged in terms of what the digital surrogate says in relationship to the source document itself, not what the technology was capable of doing. Being able to make those basic concepts relevant in a technical realm should be an important part of formal education.”

Almost everyone felt that LIS programs were important venues for basic professional education but that increasingly we also need people with backgrounds in computer science, business, law, and other related fields to deal with all aspects of digital records.



- Another informant weighed theory against practice: “The most . . . critical part of educational programs is that people learn concepts first and practices as instantiations of concepts.”
- Another said, “It is not the techniques that are particularly different, it is the way of thinking about how they are applied—which is where the change happens.”
- Almost everyone felt that LIS programs were important venues for basic professional education but that increasingly we also need people with backgrounds in computer science, business, law, and other related fields to deal with all aspects of digital records. One observed, “The cultural materials community has a ‘preservation mandate’ at the heart of its professional ethics. It has the richest, most complex, and varied conglomeration of materials that need to be preserved in either traditional or digital formats.”
- Some informants—and one in particular—emphasized that we also need cross-over skills from business, computer science, law, and engineering. “This can happen in several ways: LIS students at some schools can earn joint master’s degrees; we can recruit students into librarianship and archives who already have these backgrounds; and we can provide many continuing education opportunities for professionals.”

In the section on “Preservation Planning, Actions, Procedures, and Policies,” we first asked the informants to identify three key preservation problems that affect born-digital content. Although many of the respondents initially said something like, “the usual three,” a wide variety of answers emerged, as is shown in figure 1, Key Challenges. This diversity points out the many challenges presented by digital assets.

- One informant summarized the problems cogently when he said that there is a “lack of emphasis on management issues that will make preservation routine, systematic, and implemented at the right times for the right kinds of things.”
- The question “What might be some advantages to developing national preservation policies?” turned out to be something of a cultural litmus test. Almost to a person, Americans were skeptical of the efficacy of national policies. One, who spent several years working for the Research Libraries Group in the UK, noted “while national policy seemed to work as a driver in [the UK], Americans are not a nation prone to cultural policy making. [Instead] we have focused on guidelines.”
- Two informants indirectly alluded to this ambivalence as well: “Part of the problem reflects our uncertainty in this country [USA] about who is responsible for preservation of library resources. In other developed countries, a national library has responsibility for acquiring and preserving a country’s published output. Our Library of Congress serves as a national library in some respects but has no universal preservation charge for the nation—[even though] it has recently been assigned the responsibility . . . to support long-term preservation of digital content through the National Digital Information Infrastructure and Preservation Program (NDIIPP).”⁷⁵

Perhaps as a consequence, preservation activities in the United States have tended to be de-centralized. Instead, Americans have established programs, guidelines, and standards through professional organizations, consortia, and other cooperative activities.



- A couple of the American informants did cite some possible advantages to having a national preservation policy, such as that “it would raise the visibility and awareness of digital preservation issues” and that it could “form a standard of minimum practice” for institutions.
- The non-American respondents identified two additional advantages of having national preservation policies: there would be no preservation have-nots, and national policies could be aimed at coordination and optimization.
- They also identified some potential disadvantages. One noted that there are downsides to national policymaking and that one could “lose the leading edge . . . like the committee that designs the horse and winds up with the camel.” Another speculated that such policies might cause some programs to fold.
- In answer to “Whom do you see as the prospective future users of digital materials?” every respondent indicated that everyone is a prospective user.
- We asked about cost, in its broadest interpretation, as well as its role in decision-making. Regarding this broad view, most respondents felt that inaction may bring about inestimable loss of our cultural heritage. One informant noted: “We measure dollars and we measure production, but we don’t measure how well the program meets the needs of an institution. I get concerned when people look at cost and start making assumptions about how successful a program is. With respect to ‘cost benefit,’ you must look at it in terms of the cost of *not* doing something, the cost of doing it, and the cost saved by doing it as part of a consortial effort of contributing institutions. We must be wary of how we define ‘cost assessment’ in terms of decision-making. Cost is reality training, first and foremost.”
- And in a similar vein, another asked, rhetorically, “How do you justify putting resources into preservation of digital records, how do you demonstrate the benefit to the community, and how do you provide some kind of notion of what it would cost the community not to have that happen?”

Other perspectives emerged:

- “We are at the birthing stage of developing costs. If we do things right, we won’t need an assignable cost—it will just be a part of what we do, like mopping the floor.”
- “Cost can be an impediment to moving forward because people use cost as an excuse for not acting.”
- “A lot of the costs in the long-term maintenance issues for digital materials get back to [the costs of] continuous appraisal.”

And similarly, one respondent showed that the problem is not new or unique to digital texts:

- “Some issues are the same as in the paper world: what do you select, how do you keep it. Also, there is the cost of maintaining that information over a period of time.”

Informants said we must not merely find short-term solutions. This perspective is consistent with our survey results.



One of the two wrap-up questions was “What are your great concerns regarding the preservation of records?” The responses to this question mentioned lack of public awareness, inadequate funding, rights management issues, and social and political issues. This overlapped with our earlier question about key challenges, but here the informants were concerned with a broader social context. For example:

- “One of my concerns is the will to preserve in a disposable society. Benign neglect is not an option in a digital world.”
- “We must engage the public in such a way that they understand that preservation is a social and cultural activity.”
- “Migration makes archival appraisal a continuous process; every time you have to do a migration, you have to reappraise.”
- “One of the things we are concerned about is [sic] the records in our custody: can we certify that they are what we say they are? The chain of custody is important. One of the biggest concerns is insuring the integrity of the record.”
- “Vast amounts of money are going to preserve government records, yet no two agencies have the same practices.”

One informant described what he identified as “public good” versus “free rider” issues. “In the digital medium, where it’s very easy to copy things from one place to another, it . . . brings to the forefront the whole ‘public good’ problem that preservation represents anyway, and that is, perhaps, something the culture at large benefits from, if a single institution takes responsibility for preserving an object. . . .And yet it’s not always the case that the preserving institution can get the financial support of the public that benefits from it. So it is often a huge ‘free rider’ problem and it leads to institutions saying, ‘Why should I bother? Let someone else take care of it.’”

Finally, we asked the informants to discuss current projects. Since the projects were institution specific, the answers did not yield results we could generalize for the present study.

Conclusions

The data collected from the surveys show that since round 1:

- There is still a lack of consensus on preservation strategies and most institutions are continuing to explore one particular type.
- Institutions are beginning to address cost issues.
- An increasing number of institutions are developing policies or seeing the need to develop them.
- Respondents have a clearer idea of their staffing needs over the next three to five years than was indicated by the round 1 participants.
- Nearly half of the institutions surveyed in round 2 have secured funding for R & D—through budget increases, government appropriations, collaborations, and grants.
- Institutions are actively continuing to develop strategic alliances and collaborations that are largely centered on project-based R & D.



The key informants identified many digital preservation challenges including technical, social, political, legal, and educational issues (figure 1). They stressed the need for technical solutions such as more standards and a greater focus on such issues as interoperability, scalability, and automated metadata. At the same time, the informants lamented “the muddy waters of rights management” as well as the lack of awareness of digital preservation issues among the general public; they identified the need to educate people to address these issues; they said that ongoing, continuing education is crucial, but there were different ideas about how to educate future professionals. Some felt that with a solid liberal arts education one could learn “on the job.” Others felt that there was a need for more professional librarians and archivists as well as computer scientists, economists, lawyers, and managers. Those who felt that a master’s degree in library, archival, and information science was crucial also believed that current programs are not addressing preservation needs adequately.

Taken together, the surveys and key-informant interviews indicate that:

- Both the nuts-and-bolts institution-specific approach of the survey respondents and the global perspectives voiced by the key informants are crucial to an understanding of the full range of digital preservation issues.
- Both groups are concerned with issues relating to the authenticity and reliability of records in electronic systems. Survey respondents are trying to incorporate these concepts into their institutional procedures while the key informants are trying to create standards, models, and international initiatives that respect authenticity and reliability.
- Both groups expressed the need for more research in a variety of areas.⁷⁶

There was one unexpected outcome to the research. We did not ask the key informants about professional or theoretical influences in their work, yet everyone cited influential readings or scholars. A future strategy might be to study the impact of specific influences on the approaches that key informants bring to their work. As far as this study went, this was secondary data that emerged through data analysis, and we chose to include it as appendix 5, http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v005/supp_open/5.2cloonan_appendix05.html.

What do we need to know? What are the implications of our findings? During the three-year period of our two surveys, projects progressed or were completed, and the practitioners interested in the challenges of long-term preservation of electronic records now have results to review. Institutions that are actively involved with seeking solutions to these problems were represented in this survey. The wealth of information participants provided illustrated the rapid technological and intellectual progress made during the last three years. There is now more thoughtful literature available on problems related to long-term digital preservation than there was three years ago. Also, various camps have solidified behind particular preservation strategies. In the short time between surveys, we moved from managing (and needing to preserve) mostly flat files to dealing with files with links to multi-media, dynamic, experiential, and interactive records.

As we stated at the beginning of this study, our purpose is to identify approaches to and ideas about digital preservation. We reported the findings of round 2 of our



study, in which we continued to document extant practices as well as expert opinions on the digital preservation dilemma. We explored some known roadblocks to long-term digital preservation—such as methods, costs, intellectual property rights issues, and selection and appraisal for preservation.

We particularly note that there is not as yet a consensus on a single preservation strategy. Should we expect it or want it? We suggest that an area of further research might be an exploration of several forms of preservation strategies within one project—a suite-of-tools approach. It may be that one type of preservation strategy is more effective for particular types of records than another, or a particular type of preservation strategy may be more cost-effective to use than another for a particular institution's needs. At some point, we as a community must find answers to these challenges and take the opportunity to compare the performance of various preservation strategies and their costs across types of records—as an institution would have to do in order to make choices appropriate to its operations and mission. In an ideal world, the pragmatism of the survey respondents and the broad perspectives of the key informants should function to support each other's efforts. This may become the next step toward taking research closer to real-world applications, which, after all, is the heart of issues we have explored.

Neal Beagrie makes three observations about the institutions that he studied that complement our findings:

In none of the countries surveyed [Australia, France, the Netherlands, and the United Kingdom] is there a single national initiative for digital preservation. Rather, there are many institutional missions that are being extended to the digital domain, including those of national institutions such as national archives and national libraries. . . . Digital preservation is poorly funded in relation to the scale of the problem, and . . . it is [still] easier to obtain funding for digitization for access than for preservation. The long-term benefits and requirements of preservation seem often to be overshadowed by the immediate benefits of current access initiatives.⁷⁷

At the end of round 2, we find that preservation activities for digital content are often ad hoc or inconsistent. Yet it is promising that some of our survey respondents are thinking strategically by working with others across institutions. In round 3 we are taking a close look at strategies employed at two of the institutions surveyed. Cloonan is conducting a case study at WGBH in Boston that is focusing on the development of the Universal Preservation Format (UPF). Sanett is studying the processes for managing and preserving electronic records at the National Archives of Australia.

Michele V. Cloonan is dean, Graduate School of Library and Information Science, Simmons College, Boston, MA; she may be contacted via e-mail at: Michele.cloonan@simmons.edu.

Shelby Sanett is director of special projects, Amigos Library Services, Dallas, TX; she may be contacted via e-mail at: sanett@amigos.org.



Notes

1. Rounds 1 and 2 were prepared, in part, with support from the National Historical Publications and Records Commission (NHPRC), the National Science Foundation (NSF), the Social Sciences and Humanities Research Council of Canada (SSHRC), Amigos Library Services, Simmons College, and the State University of New York at Albany.
2. See Richard Pearce-Moses, *A Glossary of Archival and Records Terminology* (Chicago: Society of American Archivists, 2004) for the variety of definitions for these terms.
3. Anne Gilliland, e-mail to the authors, January 20, 2004.
4. The definitions are taken from the InterPARES *Glossary*, <http://www.interpares.org/book/index.cfm> (accessed January 20, 2005). It became apparent in the interviews and in the use of library-centric vocabulary that the researchers had a background stronger in library practice than in archival practice. For example, we used the term “materials” when referring to all the holdings in a library, whereas archivists use “archives” or “records” to mean the same thing. Sometimes our use of library terms temporarily confused some interviewees.
5. Thanks to Ken Thibodeau for providing us with this, his working definition, personal e-mail to Shelby Sanett, April 22, 2004.
6. For all the findings from InterPARES 1, see the InterPARES Project Web site, <http://www.InterPARES.org>. (InterPARES 2 [IP 2] is a five-year continuation project that is “applying a multi-method approach to the development of concepts, processes and tools that will help in the securing of a protected and lasting environment for the digital records produced in interactive, dynamic and experiential systems in the course of artistic, scientific, and government activities.” Round 2 overlapped with the beginning of IP 2.) See also: <http://www.InterPARES.org>, IP 2 for further information (accessed January 20, 2005).
7. The first person plural pronouns in this article refer to its two authors, as well as the collective “we”: information professionals.
8. For the full preservation model, see the final report of the Preservation Task Force, *Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, www.interpares.org/book/index.htm (accessed January 20, 2005). The model is based on the Integration Definition for Function Modeling, IDEF(0), constructed within the Reference Model for an Open Archival Information System (OAIS).
9. Neal Beagrie, *National Digital Preservation Initiatives: An Overview of Developments in Australia, France, the Netherlands, and the United Kingdom and of Related International Activity* (Washington, D.C.: Council on Library and Information Resources and the Library of Congress, April 2003).
10. *Ibid.*, 1.
11. Michele V. Cloonan and Shelby Sanett, “Preservation Strategies for Electronic Records: Where We Are Now—Obliquity and Squint?” *The American Archivist* 65 (Spring/Summer 2002): 70–106.
12. Institutional interview #3, conducted September 1, 2001, in Washington, D.C. We refer to the institutions by number, but the list of participants appears in Appendix 3 (http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v005/supp_open/5.2cloonan_appendix03.html).
13. We drew the definition of “key informants” from two sources: John W. Creswell, *Qualitative Inquiry and Research Design: Choosing Among Five Traditions*, 1st ed. (Thousand Oaks, CA: Sage, 1998) and *The American Heritage Dictionary*, 3rd ed. Creswell defines “key informants” as “individuals who provide useful insights into the group and can steer the researcher to information and contacts,” 60. *The American Heritage Dictionary* defines “informant” as “One who furnishes linguistic or cultural information to a researcher,” s.v. “informant.”
14. John W. Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 2nd ed. (Thousand Oaks, CA: Sage, 2003): 19–20.



15. Grant McCracken, *The Long Interview* (Newbury Park, CA: Sage Publications, 1988): 12.
16. Policy in the survey referred to institutional practices, whereas we asked the key informants to address national preservation policies (see Appendices 1 and 2, http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v005/supp_open/5.2cloonan_appendix01.html and http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v005/supp_open/5.2cloonan_appendix02.html).
17. The transcribers were Kevin Glick (now at Yale University), Reg White, and Mark Wolfe. Jonathan Lill, then a student at the Graduate School of Library and Information Science at Simmons College (and now at Columbia University), helped with the editing, and Francesca Marini, a doctoral candidate in the Department of Information Studies at UCLA, assisted with the transcripts and citations. We acknowledge their work with many thanks.
18. In the last group interview, the interviewer and one interviewee were in New York City, while the other, in California, participated by telephone.
19. See Creswell, *Qualitative Inquiry*, p. 202, where he defines triangulation as using multiple sources to provide corroborating evidence. In his second edition (2003), Creswell names our approach the “concurrent triangulation strategy” (p. 217). The definition is essentially the same: “two different methods in an attempt to confirm, cross-validate, or corroborate findings within a single study.”
20. For definitions, see Cloonan and Sanett, “Preservation Strategies,” 95–6.
21. Institutional Interview #5, July 30, 2001. In the subsequent text, many statements of Interview #2 corroborate what others have said.
22. Institutional Interview #4, November 14, 2001.
23. Institutional Interview #3, September 1, 2001.
24. Institutional Interview #7, October 18, 2001.
25. Institutional Interview #6, February 25, 2003. See Gregory W. Lawrence et al., *Risk Management of Digital Information: A File Format Investigation* (Washington, D.C.: The Council on Library and Information Resources, June 2000).
26. Institutional Interview #1, September 11, 2001.
27. Institutional Interview #3, September 1, 2001.
28. Institutional Interview #4, November 14, 2001.
29. Institutional Interview #6, February 25, 2003.
30. Institutional Interview #5, July 30, 2001.
31. Institutional Interview #3, September 1, 2001; Institutional Interview #5, July 30, 2001.
32. Institutional Interview #1, September 11, 2001.
33. Institutional Interview #3, September 1, 2001.
34. Institutional Interview #5, July 30, 2001.
35. Institutional Interview #7, October 18, 2001.
36. Institutional Interview #8, March 1, 2002.
37. Institutional Interview #1, September 11, 2001.
38. Institutional Interview #3, September 1, 2001.
39. Institutional Interview #5, July 30, 2001.
40. Institutional Interview #6, February 25, 2003.
41. Institutional Interview #7, October 18, 2001.
42. Institutional Interview #8, March 1, 2002.
43. Institutional Interview #1, September 11, 2001.
44. Institutional Interview #5, July 30, 2001.
45. Institutional Interview #6, February 25, 2003.
46. Institutional Interview #1, September 11, 2001.
47. Institutional Interview #5, July 30, 2001.
48. Institutional Interview #7, October 18, 2001.
49. Institutional Interview #1, September 11, 2001.
50. Institutional Interview #5, July 30, 2001.



51. Institutional Interview #7, October 18, 2001.
52. Institutional Interview #1, September 11, 2001.
53. Institutional Interview #3, September 1, 2001.
54. Institutional Interview #5, July 30, 2001.
55. Institutional Interview #6, February 25, 2003.
56. For a discussion of Moore's Law see Gordon E. Moore, "Cramming More Components onto Integrated Circuits," *Electronics* 38 (April 19, 1965), <http://www.intel.com/research/silicon/mooreslaw.htm> (click on "original paper").
57. Institutional Interview #7, October 18, 2001.
58. Two recent examples include the National Archives/ICA Conference, Kew, England, April 2003 and the ARSAG Conference, Paris, France, May 2002.
59. Institutional Interview #5, July 30, 2001.
60. Institutional Interview #5, July 30, 2001.
61. Institutional Interview #7, October 18, 2001.
62. Institutional Interview #3, September 1, 2001.
63. Institutional Interview #1, September 11, 2001.
64. Institutional Interview #3, September 1, 2001.
65. Institutional Interview #5, July 30, 2001.
66. Institutional Interview #8, March 1, 2002.
67. Institutional Interview #5, July 30, 2001.
68. Institutional Interview #3, September 1, 2001.
69. Institutional Interview #5, July 30, 2001.
70. Institutional Interview #6, February 25, 2003.
71. Lynch, "Check Out the New Library," *Ubiquity* 4, 23 (July 30–August 5, 2003), http://search.netscape.com/ns/boomframe.jsp?query=Clifford+Lynch+Ubiquity&page=1&offset=0&result_url=redir%3Fsrc%3Dwebsearch%26requestId%3D6f00dedccef6bd25%26clickedItemRank%3D1%26userQuery%3DClifford%2BLynch%2BUbiquity%26clickedItemURN%3Dhttp%253A%252F%252Fwww.acm.org%252Fubiquity%252Finterviews%252Fpf%252Fclynch_1.html%26invocationType%3D-26fromPage%3DNSCPTop%26amp%3BampTest%3D1&remove_url=http%3A%2F%2Fwww.acm.org%2Fubiquity%2Finterviews%2Fpf%2Fclynch_1.html (accessed January 11, 2005).
72. Michael Lesk, *Understanding Digital Libraries* (San Francisco: Morgan Kaufmann, 2004), 291.
73. As these quotes were taken directly from the interview transcripts, they are not individually footnoted.
74. Cloonan and Sanett, 84–7.
75. Deanna Marcum and Amy Friedlander, "Keepers of the Crumbling Culture: What Digital Preservation Can Learn from Library History," *D-Lib Magazine* 9, 5 (May 2003), <http://www.dlib.org/dlib/may03/friedlander/05friedlander.html>. Some of the points that Marcum made in our interview with her also appear in this article.
76. A recent paper does an excellent job of identifying areas that researchers need to address. See Margaret Hedstrom and Seamus Ross, *Invest to Save: Report and Recommendations of the NSF-DELOS Working Group on Digital Archiving and Preservation*, prepared for the National Science Foundation's Digital Library Initiative and the European Union under the Fifth Framework Programme by the Network of Excellence for Digital Libraries (DELOS), 2003.
77. Beagrie, 2.