



# InterPARES Project

International Research on Permanent Authentic Records in Electronic Systems

---

---

## Survey of Preservation Practices and Plans

---

---

### Draft Appendix

*June 1, 2001*

Michèle V. Cloonan and Shelby Sanett

Preservation Strategies for Electronic Records, Round 1 (2000-2001)  
*Where We Are Now: Obliquity and Squint?*  
A Report to the National Historical Publications and Records Commission  
Michèle V. Cloonan and Shelby Sanett  
University of California, Los Angeles  
June 1, 2001

## Introduction

Information has never before been as fugitive as it is today. Whereas records were once written on media which could last hundreds—or even thousands—of years, electronic records are in danger of disappearing, becoming physically unusable or legally inadmissible, almost immediately. There are many causes for the short life span of digital and electronic records: media deterioration, technological obsolescence, a paucity of standards and guidelines, and the failure of many managers to plan for the maintenance and preservation of electronic records. Besser refers to this last cause as “the custodial problem.”<sup>vi</sup>

Records created or maintained in electronic form, either analog or digital, are herein referred to as *electronic records*. Electronic records are the focus of the InterPARES Project. It is impossible to avoid using the term *digital preservation*, however, because that is the term most frequently used in the preservation literature. (See Appendix 4, Preservation Bibliography. Nearly every author cited in this report uses it.) We recognize that it is not always the most precise term for archivists because it refers to the preservation of reformatted items, born-digital electronic materials, and born-again digital materials. We will use the term *digital preservation* to reflect the general preservation discourse found both in the literature and in our interviews, as well as *preservation of electronic records*, when appropriate. We have provided a glossary below of terms used in this study. Recently, the term *long-term retention* seems to be gaining currency.<sup>ii</sup> As this term is more general than either *digital preservation* or *preservation of electronic records*, we will also use it where appropriate. We believe that the flux surrounding terminology reflects an evolving new field.

It is the responsibility of archivists to ensure the authenticity, reliability, and long-term accessibility of permanent electronic records.<sup>iii</sup> Traditionally archivists have done so by gathering documents, establishing provenance, and maintaining and demonstrating an unbroken chain of custody in an evidence-based approach to managing records.<sup>iv</sup> Is it possible to ensure authenticity and reliability of records regardless of their formats? If so, which technologies and techniques are archivists using to preserve electronic records at the time of their creation as well as throughout their life cycle? And how are archivists defining the term *preservation*?

## Purpose

The purpose of this study is to identify and describe the strategies and techniques that are in use or in development, in institutions and on research projects, to preserve electronic records. Such knowledge should inform archivists as they formulate strategies, policies, and standards for preservation.

Preservation of electronic records is a young field. Indeed, some people believe that the very concept of electronic record preservation is an oxymoron. The term *preservation* implies permanence, yet such media are inherently unstable. Add to the ongoing problem of technological obsolescence such challenges as copyright and fair use; organizational and custodial issues; different stakeholders and interests; substantial financial requirements and legal admissibility requirements; and prospects for preservation of electronic records seem grim, indeed.

It is possible to develop models, protocols, and standards that factor in such problems as media deterioration and technological obsolescence, and work is already being done in these areas. For example, some institutions already safeguard their digital files. Others maintain metadata separately from their master files of electronic records. The implementation of more standards may satisfy the basic requirement for preservation: to maintain authentic and reliable records for as long as they are needed.

Yet as Gilliland-Swetland has pointed out,

Counterintuitively, perhaps, it is during the preservation of digital materials that evidential value is often most at risk of being compromised. Digital preservation techniques have moved beyond a concern for the longevity of digital media to a concern for the preservation of the information stored in those media during recurrent migration to new software and hardware. In the process, many of the intrinsic characteristics of information objects can disappear—data structures can be modified and presentation of the object on a computer screen can be altered.<sup>v</sup>

Since the field of electronic record preservation is fledgling, with many concepts and issues still to be fully comprehended and resolved, we decided to conduct this study in three rounds, from 2000 through 2003. Round 1 (2000-2001) surveyed 13 institutions, programs, and projects. This round also provides the baseline for rounds 2 and 3. In round 2 (2001-2002) we are administering a revised version of the questionnaire, interviewing key informants, and conducting case studies for two of the respondents surveyed in round 1. The third round (2002-2003) will focus on an even smaller number of programs and projects. By the end of the three rounds, we expect to present a sharper picture of preservation strategies in practice than we currently have. These strategies include preservation techniques (e.g., refreshing, migration, emulation); selection for preservation; staffing configurations; cost modeling; access to preserved records; and policymaking.

Although other studies have explored individual aspects of electronic and digital preservation—such as standards, intellectual property, or specific techniques such as emulation—we are examining the continuum of activities that constitute the broad range and long-term goals of preservation of electronic records. By documenting the variety of approaches that are being taken, we hope not only to shed light on current practices, but by the end of round 3, to offer informed consideration on where preservation might be headed.

When we began our research, no study had yet provided a comparison of preservation techniques for electronic records across institutions and projects. The closest research to ours is a study that was carried out by Hedstrom and Montgomery (1998) to examine long-term retention needs and

requirements in the Research Libraries Group (RLG) member institutions.<sup>vi</sup> Another related work is a book edited by Kenney and Rieger on digital imaging for libraries and archives.<sup>vii</sup> One section of the book, “Digital Preservation Strategies,” gives a succinct overview of techniques such as refreshing, migration, emulation, etc. We designed this study to examine specific long-term retention strategies in more detail.

We undertook this research as part of our work on the InterPARES Preservation Task Force in order to explore one of the five research questions that the Task Force was charged with examining: What methods, procedures, and rules of long-term preservation are in use or being developed? Our three-round study seeks to ascertain the strategies that are either currently in use or are in development for preserving electronic records. We also plan to track the evaluative studies of these techniques. (For a complete list of the research questions, see Appendix 5.)

## Research Questions

1. 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?

In addition to question 1, formulated by the InterPARES Preservation Task Force, we asked three additional questions as a result of responses to the questionnaire:

2. What is the meaning of *preservation*?
  - a. Does the meaning change when it is applied to electronic rather than paper-based records?
3. Will current strategies for preserving electronic records ensure longevity and authenticity?
4. How are costs for the preservation of electronic records derived? Have effective cost models been developed?

## Definitions

We begin this section with the two broadest terms: preservation and digital preservation. From there, we define digital preservation techniques. These relationships are hierarchical, with

preservation at the top, followed by digital preservation and digital preservation techniques. The preservation of electronic records falls under digital preservation.

Preservation

Digital Preservation

Electronic Records Preservation

[Digital Preservation Techniques]

Refreshing

Migration

Knowledge-Based Persistent Object Preservation

Emulation

Bundling

Universal Preservation Format

[Other Techniques]

Robotics (a capture, not preservation technique)

Digital Archaeology (“*post hoc* rescue”)

**Preservation:** the processes and activities which stabilize and protect cultural heritage materials so that they will be permanent and durable, or as long-lasting as it is possible to make them (authors).

**Digital Preservation:** the processes and activities which stabilize and protect reformatted and “born digital” authentic electronic materials in forms which are retrievable, readable, and usable over time (authors).

### Digital Preservation Techniques

**Refreshing:** periodically moving records from one storage medium to another. It is a preventive measure and, because of rapid media obsolescence, it will be a necessary strategy for some years to come (authors).

**Migration:** the process of moving records from one hardware and/or software platform to another (authors).

**Collection-Based Persistent Object Preservation (aka, Knowledge-Based Persistent Object Preservation):** “. . .the retrieval or instantiation of the collection onto new technology. . . . [t]he processes used to ingest a collection, transform it into an infrastructure independent form, and store the collection in an archive comprise the persistent storage steps of a persistent archive. The processes used to recreate the collection on new technology, optimize the database, and recreate the user interface comprise the retrieval steps of a persistent archive. The two phases form a cycle that can be used for migrating data collections onto new infrastructure as technology evolves” (Moore, et al., April 2000, <http://www.dlib.org/dlib/april00/moore/04moore-pt.2>).<sup>viii</sup>

**Emulation:** an applications software approach that recreates the technical environment required to view earlier programs. Such software can theoretically mimic every type of application ever written and be run on current computers (authors).<sup>ix</sup>

**Technology Preservation:** “preserving the technical environment that runs the system, including software and hardware such as operating systems, original application software, media drives, etc. While technology preservation means preserving the technical environment rather than reengineering it, as emulation does, many of the same issues apply to both” (Kenney/Rieger).

**Bundling:** “...taking objects such as Word documents and, by using software, creating bundles of documents on an independent platform. This is another form of emulation” (PRO interview; permission not yet received).

**Universal Preservation Format (UPF):** “Deals with multiple formats. It takes a snapshot of the information and gives you everything you need to view it. It is difficult to capture the whole content, e.g., interactivity, without the original equipment. Closer to emulation than migration” (WGBH interview; permission not yet received).

**Digital Archaeology:** “...accessing digital materials where the media has become damaged (through disaster or age) or where the hardware or software is either no longer available or unknown. The authors refer to this as “*post hoc rescue*” (Ross/Gow).<sup>x</sup>

**Robotics:** the use of robots to download electronic documents. Downloading in and of itself does not preserve the records (authors).

## Terms

**Record:** a document made or received and set aside in the course of a practical activity (InterPARES Glossary).

**Authentic record:** a record that is what it purports to be and is free from tampering or corruption (InterPARES Glossary).

**Digital record:** a record that now exists in electronic form though it may or may not have been created in electronic form. For example, a *digital record* may have been created on paper and digitized later. Subsequent digitization may remove or deplete its “recordness” (authors, Gilliland-Swetland).

Electronic record: **A record that is created (made or received and set aside) in electronic form (InterPARES Glossary).**

**Program: an ongoing set of services, around a common goal or activity, usually located within a single institution (authors).**

**Project: a specific undertaking or research endeavor, usually with special funding. Projects may take place within single, institutional programs, or at more than one site. Projects usually take place within a finite period (authors).**

## **Research Methodology**

To answer our overarching questions, we decided to use a purposive sampling strategy—one that would show different perspectives on the problems we wanted to address—of archives, projects, and programs in the United States, Canada, and Europe. We chose to collect data using a questionnaire which we developed with feedback from other members of the InterPARES Preservation Task Force.

We identified 15 sites that we knew were developing one or more of the following techniques: refreshing, migration, emulation, collection-based persistent object preservation, bundling, and robotics. We ultimately interviewed representatives from 13 of the 15 sites that we selected (Appendix 1, “List of Sites”).

Such a small population did not warrant a quantitative research design. Further, since we knew that we would be asking broad, sometimes open-ended questions, we decided on a qualitative design. Since it was not our goal to make statistical inferences, but to learn about processes and methodologies, we explored one research method and one research strategy: the case study and the survey.

We ruled out the case study method for round 1 because we felt that it was too early in the development of long-term retention strategies to study individual programs in depth. (This is a methodology we plan to use in rounds 2 and 3.) Rather, using a questionnaire for round 1, we hoped to establish a benchmark of current practices from which we could collect a general or comprehensive view. The questionnaire would provide baseline data about the current (year 2000) state of long-term retention practices that we could draw on in rounds 2 and 3 of the research.

With respect to procedures followed during the data-collection process, the participants were sent a consent letter which explained that if they volunteered to participate in this study, they agreed to read over the attached survey instrument (See Appendix 2, Questionnaire) and participate in a telephone interview based on this instrument. We followed-up on this letter by telephone or e-mail to arrange a suitable time for the interview. The telephone interview was not taped. In a few instances we were able to conduct the interviews in person. To the best of our knowledge, there were no foreseeable risks, discomforts, or inconveniences to anyone who participated in this study. In terms of the information obtained from this study, we agreed not to disseminate proprietary information, quote any of the interviewees, or disclose individual or institutional identities without express written permission.

Volunteers had the option to withdraw from the study at any time without consequences of any kind. They could also refuse to answer any questions they did not want to answer and still remain in the study. As investigators, we also maintained the option of withdrawing participants from the research if circumstances arose which warranted doing so.

### **Boundaries and Limitations of the Study**

The data gathered and analyzed in round 1 of this research allow us to draw only tentative conclusions about current preservation techniques, and in a limited number of venues (6 archives, 6 programs and research projects, and 1 library). Although our aim was to identify as many different preservation techniques as possible—without regard to how many institutions and projects were experimenting with new techniques—there is the possibility that we missed learning about important new projects. Also, preservation techniques such as bundling and Universal Preservation Format (UPF) are still only in the earliest stages of development, so we did not learn as much about them as we had hoped. However, the study has the advantage of identifying and describing cutting-edge approaches to the preservation of electronic records and the researchers will continue to monitor new developments.

Another limitation is that InterPARES sponsored this research, and many of the participants in our study were affiliated with the project. Therefore, the needs and perspectives of the InterPARES Preservation Task Force drove the study. Since InterPARES focuses on electronic records, we focused on archives and archival projects. We selected one library, but the practices of that library demonstrated it to be a disconfirming case and we will not include it in round 2.

On the other hand, InterPARES is an international project with great visibility. Our affiliation with the project may have resulted in entrées to research projects that we might not have otherwise been able to gain access to.

The responses from our participants were probably not as comprehensive as they might have been had we taped them. We chose not to tape the interviewees thinking that it would cause the participants, as representatives of their institutions, programs, or projects, to be constrained in their responses. During the interviews we both took notes. After the interviews were completed, and upon further review of our notes and procedures, we felt that our notes represented more of an interpretive record of the interview than a record of quotations. The participants did not appear to be constrained in their responses, which led us to consider the possibility of taping subsequent interviews. We informally surveyed some of the participants and learned that they do not object to having their interviews taped for the second round of the survey. We feel that using the telephone during the interviews also introduced a barrier between the respondents and us. This was especially apparent in interviews where respondents spoke English as a second or third language. In the next round of the preservation survey, we have decided to ask for electronic responses to the survey instrument, and to tape the follow-up telephone interviews. When possible, we will conduct in-person interviews.

A final potential limitation of this study relates to concepts and terminology, a limitation pointed out by Hedstrom and Montgomery in their RLG study.<sup>xi</sup> We also surveyed practitioners and

researchers in both the United States and abroad. Like their study, we found that differences in terminology and concepts may reflect different cultural perspectives, or, since our interviewees included archivists, librarians, computer scientists, and engineers, professional differences.

## **Results**

The questionnaire was divided into 14 sections:

- A. Information about the Institution/Project and Respondent
- B. Program and Policy
- C. Specifics of Preservation Technique/Method/Strategy
- D. Selection for Preservation
- E. Cooperation
- F. Staffing
- G. Technical Questions
- H. Costs
- I. Preserving Records
- J. Description/Documentation of Preservation Processes
- K. Access to Preserved Records
- L. Charges
- M. Reproduction and Copyright
- N. Preservation Policies

Not all sections or questions pertained to each project or program. However, we chose to be comprehensive in order to learn about as many aspects of each program as possible. We will revise the questionnaire in the 2001 iteration of the survey. In this iteration we tried to achieve breadth and scope; in the next version we will focus on some of the above sections in greater depth. (See Appendix 3, Tabulated Responses.)

We have followed the procedures for Human Subjects research established by the University of California. Respondents were guaranteed anonymity unless they provided express written consent. Although we have requested such consents, at the time of this writing we have not yet heard from all of the respondents. All quotations used here will be designated as respondent 1, 2, etc., and references to projects as institution 1, 2, etc.

### **Section A: Information about the Institution/Project and Respondent**

This section is summarized in Appendix 1.

### **Section B: Program and Policy**

The questions were:

3. Please describe your institution's program or activities related to preserving digital objects over long periods of time.

- 3.1. When did your institution's program or activities begin?
- 3.2. Describe any institutional issues which impact upon the program.
  - 3.2.1. Describe in broad terms what methods or techniques you are exploring or using for digital preservation.
- 3.3. Describe the digital materials your institution is preserving.
  - 3.3.1. Do you consider any of these materials to be records?
- 3.4. Do you make any special provisions for preserving records, as opposed to other types of digital materials? If so, what?
- 3.5. Has the program or activity reached the point of either testing or evaluating any of the methods or techniques you are using? If so, what are the results to date?
- 3.6. Have you identified any problems, difficulties or threats to the integrity of the digital materials resulting from the use of these methods or techniques? If so, please describe.
- 3.7. How do you use the word "preservation" at your institution? In other words, what definition does your institution associate with the term "preservation"?

In this section we tried to ascertain the range of preservation activities for electronic records. We started by asking each respondent to give an overview of his or her program. When asked, "When did your institution's program or activities begin?" respondents traced electronic preservation programs back to the inception of *any* preservation activity in their institutions. Four respondents reported that their programs started in the 1970s, 1 in the 1980s, and 5 in the 1990s. Of the 4 who reported activity since the 1970s, 2 archives reported that they had been preserving digital objects since the 1970s. Three responded "not applicable" because they represented special projects. Each respondent then described the development of preservation programs over time.

Respondents were asked to describe institutional issues that impact their programs. Issues included: the problems associated with outsourcing; inadequate staffing; storage; prioritization of records to be preserved; inadequate resources, including funding; and increasing legal mandates for preservation.

We also asked what methods or techniques they were exploring or using for digital preservation. The response: Migration=4; Emulation=2; Knowledge-Based Persistent Object Preservation=3; Bundling=1; Refreshing=1; Digital Archaeology=1; Preservation copying=1; Physical Preservation=2; and Robotics=1. It is important to note that some institutions and projects are using more than one strategy, and that Bundling is in the exploratory stages only.

The next question asked respondents to describe the items that they are preserving. Respondent 1 reported preserving only born-electronic records. Others preserved a variety of born-electronic records and other digital artifacts including spreadsheets, databases, computer games, geomatics, serials, etc.

One question in this section caused confusion: “Do you consider any of these materials to be records?” Six answered “yes,” one answered “no” and 6 answered “not applicable.” Not all of the respondents distinguished between records, documents, and items. Our follow-up question, “Do you make any special provisions for preserving records, as opposed to other types of digital materials? Only respondent 11 said “yes.” Five said “no” and 7 responded “not applicable,” perhaps due to the same confusion demonstrated in the preceding question.

We also asked if the program or project was testing or evaluating any of the methods or techniques currently being used. Only three respondents are engaged in testing or evaluation.

The next question asked respondents to identify problems, difficulties, or threats to the integrity of digital materials with any of the above-mentioned techniques. Three respondents identified preserving the integrity of the original digital object as a threat. Respondent 4 discussed the issue of “acceptable loss” due to migration. Two other respondents identified the problem of changing standards. Finally, respondent 12 mentioned the issue of migration technology obsolescence.

From this section of the questionnaire, it became apparent that the thirteen institutions and projects were better versed in identifying problems than in developing solutions.

The final question asked respondents to define preservation. Their responses are described in detail below.

### **Section C: Specifics of Preservation Technique/Method/Strategy**

4. What preservation technique does your program use?
  - 4.1. How was this method selected?
    - 4.1.1 Is it a hybrid, e.g. a combination of two or more preservation techniques/methods, such as microfilming and scanning?  YES  NO
  - 4.2. If you are using a hybrid model, how did it evolve?
  - 4.3. From other methods you have tried before or you are aware of that other repositories are using, how is this method different from other methods?
  - 4.4. In selecting the preservation method or strategy, have you considered what its effect might be upon the intellectual integrity (e.g., authenticity and reliability) of the digital material?  YES  NO

- 4.4.1. If yes, are you able to prove/demonstrate that the intellectual integrity of the digital material has not been compromised through the preservation process? Please explain.
- 4.5. Is there evaluative data on the efficacy of this preservation method/model? Please describe.

This section asked respondents to elaborate on the preservation techniques that they identified in Section B. As reported above, the most common strategy is migration. We asked respondents in this section to tell us how they selected the method. Respondent 1 identified migration as the standard method of moving from one platform to another. Respondent 7 responded that her institution believes it to be the best method.

As a follow up question, we asked if any of the institutions or projects used a hybrid approach that combined two or more techniques. Four respondents said “yes,” four said “not applicable,” one said “maybe,” and three said “no.” Of the four yes answers, three indicated microfilming and scanning and one said scanning and physical treatment. No one reported combining such techniques as robotics and migration, for example.

We also asked whether the respondents had considered what effect their chosen techniques might have on the intellectual integrity of the digital materials. Nine respondents indicated “yes;” four, “no.” Respondent 7 criticized the work of another institution that enhanced original photos, citing issues of veracity. Respondent 8 worried that emulation may not work for some classes of materials.

The final question in this section asked whether there is evaluative data on the efficacy of the preservation method/model used by each respondent. Only two respondents had formal evaluation procedures. Respondent 1 reported that audit trails are kept, and that his institution is moving to a document management program (ERM) to support audit trail development. Respondent 11 uses Knowledge-Based Persistent Object Preservation which conforms to the OAIS model.

#### **Section D: Selection for Preservation**

5. Which of these criteria guide selection of materials for preservation? (Check all relevant)
- |  |   |
|--|---|
| <input type="checkbox"/> Historical/cultural value     | <input type="checkbox"/> Save space                           |
| <input type="checkbox"/> Legal requirement to preserve | <input type="checkbox"/> Research into preservation processes |
| <input type="checkbox"/> Retard deterioration          | <input type="checkbox"/> Commercial use                       |
| <input type="checkbox"/> Increase access               | <input type="checkbox"/> Other reasons (please specify)       |
- 5.1. Where did the materials you selected come from?
- |   |  |  |
|---|--|--|
| <input type="checkbox"/> Parent institution         | <input type="checkbox"/> Government agencies                 | <input type="checkbox"/> Other (specify) |
| <input type="checkbox"/> Collaborating institutions | <input type="checkbox"/> Other organizations or associations |  |

There were 13 responses; of these there was one ‘not applicable.’ *Historical/cultural value* and *Legal requirement to preserve* were the two criteria guiding selection of materials for

preservation for 9 of the respondents. The next two criteria were *research into preservation* and *retarding deterioration* (5). *Saving space* and *Other* ranked third. No one cited commercial reasons as criteria for selection. In the *Other* category, criteria includes: materials publishers wanted them to have; institutional requirements; asset management considerations; sampling Internet sites; and supporting the curriculum. The ‘not applicable’ response is from a respondent which preserves materials for clients.

The materials which were selected to be preserved, came from: government agencies (6); parent institutions (5); other, e.g. private, commercial publishers, politician’s private papers, private individuals, commercial entities (3); and other organizations or associations, e.g. corporations (private sector) (2).

### Section E: Cooperation

6. Did you cooperate with other organizations to develop your program?

YES       NO

6.1. If so, which? (Check all relevant)

Archives       Libraries       Public companies       Museums

6.2. Is your cooperation       National       International       Local

Shared facilities       By institutional type  
[churches, labor unions, etc.]

6.3. How is the work distributed?

Equally       Work distributed in a different way?

6.3.1. Please describe.

6.4. If your program is collaborative, how did it evolve?

6.4.1. Please describe the strengths and weaknesses of the collaboration.

Of the 13 respondents, 12 cooperate with other organizations to develop their preservation program, while one does not. The types of collaborating organizations include archives (9), libraries (9), public companies (5), museums (3); and in the ‘other’ category, government agencies, other programs and projects, for profit and non-profit educational institutions; and universities. Cooperation occurred on an international level (11), national level (9), local (2) and with shared facilities (1). Work was distributed both equally (5); and in a different way (6).

Collaborations tended to evolve rather than to be planned. Participants viewed the strengths of collaboration as being shared responsibilities and costs; shared information and resources; and opportunities to develop a consensual outlook. Weaknesses of the collaborative process include resource costs, people being far apart and influenced by local interests and resources; span of control issues; and research and development going off on tangents which were non-productive.

## Section F: Staffing

7. Who is involved with the program and in what capacity (ies)?

7.1. Describe their duties.

The majority of staff is part time, work under the supervision of a person who usually holds at least one Master's degree; and divide their time with other projects and departments. Ph.D.s accounted for only two of the positions, and both worked on research projects. Educational backgrounds and skills include computer scientists, archivists, librarians, people with management and history backgrounds, and people who learned their technical skills on the job. One of the Ph.D.s is in History and Preservation; that person directs a facility.

## Section G: Technical Questions

8. Is preservation carried out by  the institution (in-house)  
 commercial vendor/contractor

8.1. Describe any pre-preservation preparation of records.

8.2. What do you consider to be the strong points of your institution's preservation methods or techniques?

8.2.1. What do you consider to be the weak points of your institution's preservation methods or techniques?

8.3. What quality control methods are applied to the preservation process or activity?

8.4. How are you storing the electronic records that have been preserved?

Ten respondents carry out preservation in-house, while (1) uses a commercial vendor. One respondent uses both in-house and commercial resources, while another respondent has disks on a shelf and is not taking pro-active steps to preserve them at this time.

Pre-preservation preparation of records includes documenting provenance; checking to make sure that records have not been tampered with; inspecting them; physical preparation as needed; putting the records into a standard format; accessioning; and arranging and describing the records prior to copying them.

The questionnaire asked respondents to discuss weak points in their institution's preservation methods or techniques. These include database problems; and some respondents working with GIS and CAD materials, found that tabular displays don't work. For electronic documents, respondents cited compound records, website material, attachments and nested materials, as being problems. Implementing finding aids for a million collections is a problem in one archive. Another archivist said that his archive has not yet done the right thing with respect to textual

documents. According to him, there are not enough resources, e.g., money and people; and the current staff does not have sufficient expertise for the work. Lastly, an archivist of a large national archive stated that the archival profession has not yet articulated its needs regarding system requirements; that the profession “is not used to system thinking in work, and never had an opportunity to do it before.” During round 2, we will explore the scope of these issues in more detail to determine whether they are ongoing, and whether resolutions have been investigated or applied.

Quality control methods applied to the preservation process or activity include the use of a quality control manual for a large European archive, with more information in development. One respondent stated that his facility was currently using *ad hoc* methods not worthy of sharing. Another respondent cited quality control of metadata as being the most difficult technical challenge. One respondent stated that quality control methods would be a partial outcome of their project. An archivist in a European archive stated that there was no really organized quality control; they rely on the professionalism of the individual, who often has no formal training. Other responses include: the use of standard information technology techniques for verifying the quality or success of the copying; and peer review research to promote publishing and discussion.

Permanent storage of electronic records includes archival storage at partner sites; containers; server/redundant servers; and underground storage. One site is not storing backup copies of records offsite.

## **Section H: Costs**

9. What do you estimate are the costs to preserve the records? (Please include staff, equipment, space, energy and other related costs)
- 9.1. If applicable, please describe the categories of your preservation costs.
- 9.2. What are the sources of funding for the program, and how are they allocated?

See below for discussion.

## **Section I: Preserving Records**

10. How are preserved records organized?
- 10.1. How is provenance respected?
- 10.2. Are there any restrictions on access to the records? If so, how are they enforced?

The majority of respondents, who answered the questions, stated that in their institutions, they organized records according to the standards currently in use, e.g., the same as for paper records; or according to the archival arrangement and description standards adopted by their institution.

Five respondents stated that the question did not apply, while one institution responded that an organizational scheme would be an outcome of their project.

Regarding respecting provenance, four respondents stated that the question does not apply, while the balance of the respondents cited the development of metadata standards, describing records according to provenance and type of record, and conformance with the institution's general practice as examples.

Six respondents stated that the question did not apply. However, all of the remaining respondents cited examples of the restrictions that their institution imposed on records. Respondent one stated that there was a 30 year closure [from general public access] unless the record was in the public domain; and that the record can in fact, be closed for up to 100 years. Respondent four also cited the 30 year closure, with the addition that the records were still subject to subpoena. Other respondents stated that the access to records is negotiated individually with each donor, while respondent eleven explained that the type of restriction on the record would determine accessibility.

## **Section J: Description/Documentation of Preservation Processes**

11. Describe record-keeping for the preserved material.

11.1. Are preserved materials described according to a recognized standard?  YES  NO

11.1.1. If yes, which one?

11.2. Are the records for preserved materials and the original  the same  
 independent of each other

11.3. How is metadata used to describe preserved materials?

Respondents generally described record keeping for the preserved material as detailed audit trails, catalogs, on databases, as part of corporate descriptive systems, or as part of the metadata which accompanied the preserved objects. The question did not apply to six respondents, because they are not yet using metadata. We expect to see more widespread use of preservation metadata in round 3. (RLG and OCLC collaborate on the international Preservation Metadata Working Group. The group is trying to establish approaches for preservation metadata that will work in a variety of settings for a variety of materials.)

Nine institutions use recognized standards to describe preserved materials. These standards include ISAD-G, EAD, MARC, modified LCSH and Dublin Core.

We asked whether the records for preserved materials and the original were the same or independent of each other and most of our respondents did not understand the question. Eight respondents said the question did not apply, while two respondents stated that the record(s) were both the same and independent. Two respondents stated that the record was the same, while

respondent ten stated that the institution does not recognize that preserved materials are copies; both copies are the records.

According to the respondents, five institutions are addressing metadata use. The question did not apply to four other institutions. Respondent six stated that the use of metadata is extensive, e.g., content description, preservation history, but that it was not always completed. We take that to mean that the metadata record was not always produced in its entirety, according to institutional standards. Respondent ten stated that metadata is being used in a variety of ways.

### **Section K: Access to Preserved Records**

12. Are the preserved records available  only on-site  only within the institution  
 through a website  other (specify)

12.1. If available through a website, please give the URL

---

12.2. Is the archival workstation equipped with  access control mechanisms  
 billing software

12.3. Is the archival workstation connected to  internal institutional servers  
 an Intranet  the Internet

The majority of institutions make records available through a website (7). One institution makes hard copy available, while one institution makes copyright protected material available to the academic community only. We did not ask whether the records were full text, or all or some series. The question did not apply to four institutions.

Question 12.2. did not apply to twelve institutions, while the remaining respondent stated that the institution had dedicated machines which consisted of five workstations, some online, and some connected to secure servers.

The last question in this section did not apply to any of the respondents.

### **Section L: Charges**

13. Do users have to pay to use the preserved material?  YES  NO

13.1. If yes,  on-site  outside the Institution  
 when accessed through the website  other (specify)

13.2. If charges are made, how are these calculated?  
 single charge  by time  by volume of material  by intended use  
(commercial/academic/students)  customized service (please describe)

other (give details)

- 13.3. If charges are made, how are they collected?  invoice  cash at point of use  
 credit card  electronic accounting  other (give details)

Seven institutions do not charge users to use preserved materials. Four institutions charge a fee for copies, two institutions charge users a fee to use preserved materials— when the material is accessed through the website. Charges are calculated by time (1), by volume of material (2), by intended use (1), e.g., commercial, academic, students, or as an institutional charge (1). One institution collects charges by credit card; three send invoices; and one is in the process of determining the procedure. One institution collects charges both by credit card and by invoice. Several respondents did not know how the charges were collected in their institutions.

### **Section M: Reproduction and Copyright**

14. Do you preserve material in copyright?  YES  NO

- 14.1. If yes, is this done  under legal provisions for your institution  with the owner's agreement  by paying the owner a fee  under license  without formalities

- 14.2. Does the institution own the copyright for the electronic form of the records?  
 YES  NO

- 14.2.1. If no, who does?

- 14.3. Are users allowed to do any of the following?  
 make printouts  download to a PC  download to a local network (LAN)  
 download to a general network (WAN)

- 14.4. Are any electronic management systems used to control copying?  YES  NO

- 14.4.1. If yes, which ones?

Ten institutions preserve material in copyright, while one institution does not. The question did not apply to two institutions. Of the institutions which preserve material in copyright, (8) stated that this was done under legal provisions for their institution, with the owner's agreement (10) and without formalities (1).

Six institutions do not own the copyright for the electronic form of the records, while the question did not apply to seven respondents. Ownership of the copyright resides in the creator, owner, or government, depending on the institution.

Users are permitted to download material to a PC (2); download to a local network (LAN) (2); or download to a general network (WAN) (2). One respondent said that the institution does not permit users to download material. The question did not apply to nine respondents.

With respect to electronic management systems in use to control copying, one respondent stated that that was being researched, while four respondents stated that their institutions do not have electronic management systems of that type in use. The question did not apply to eight respondents.

## **Section N: Preservation Policies**

15. Do you have a general preservation policy that includes records in electronic form?  
 YES       NO

15.1. If not, do you have a policy for reformatting, refreshing, migrating, emulating, or bundling data to newer technological platforms?       YES       NO

15.2. Please describe any policies you might have that relate to preservation of electronic records.

Five respondents do not have a general preservation policy in place that includes electronic records, while three respondents stated that their institutions did have such a policy in place. Two respondents indicated that their institutions had such a policy in development or being researched, while three respondents said that the question did not apply.

Regarding question 15.1., one institution has a policy in place that includes various types of preservation techniques, while four institutions do not. Respondent 10 stated that the institution has no formal policy, but that the established practice is to re-format to standard non-proprietary formats. Two respondents stated that their institutions did not have policies finalized as yet.

See below for further discussion.

Three questions in particular yielded stimulating and--we believe--pertinent observations: definitions of preservation (in the *Program and Policy* section), categories of costs (*Costs*), and preservation policies (*Preservation Policies*).

### *Definitions of Preservation*

In the section on *Program and Policy*, we asked: “How do you use the word ‘preservation’ at your institution?” In other words, what definition does your institution associate with the term ‘preservation’?” We included this question because we felt that such definitions might have a bearing on approaches that projects take to developing long-term retention strategies. For example, depending on your view of preservation you might select one approach over another. Since we interviewed archivists and librarians from six countries representing a dozen projects or

institutions, we anticipated getting a range of responses. As we continue this study into the next phase of this research, we will try to determine whether or not the definitions of preservation continue to evolve.

Eleven respondents defined ‘preservation’<sup>xii</sup>: The following key phrases emerged:

Respondent 1: “Preservation for paper records is a regime which tries to slow entropy and avoid degradation. For digital records, it is to preserve the document to perpetuity. Digital Preservation includes issues of authenticity.”

Respondent 2: “Preservation means ensuring the object is accessible over the long-term. Access and preservation are separate.”

Respondent 4: “Preservation covers all activities directed towards ensuring the ongoing accessibility to the information content of the records. Hence, we consider the ambient conditions in our repositories as a preservation issue, along with the specifications of the media on which recorded information is stored. Migration of digital objects is thus a preservation strategy.”

Respondent 5: “The ability to discover, access, and present electronic records through arbitrary changes of technology. We can preserve things forever.”

Respondent 7: “Forward migration or prospective preservation to whatever new technologies exist. [We are beyond] thinking about ‘x’ number of years of preservation.”

Respondent 8: “Enabling long-term access to materials.”

Respondent 9: “Ability to present the record unchanged repeatedly.”

Respondent 11: “Everything you have to do to guarantee you can deliver records [and] respecting the sanctity of the original order.”

Respondent 12: [The technical and managerial processes that protect the integrity and longevity of materials – regardless of genre.]

Respondent 13: “Making collections useful to scholars in the future.”

The respondents’ key phrases fall into three components of preservation: 1. preservation processes; 2. length of time for retention; and, 3. preservation outcomes.

Overall, the responses demonstrate a shift taking place from defining preservation as a once-and-forever approach for paper-based materials, to an all-the-time approach for digital materials. (Paper-based materials also require all-the-time care, mainly through environmental controls and proper storage. However, digital materials require constant refreshing, migrating, etc., a much more pro-active and costly endeavor.)

To contextualize the respondents' definitions, we offer published definitions of preservation culled from the archives and the library fields.

## **The Paper-Based Perspective**

SAA, 1974

A. The basic responsibility to provide adequate facilities for the protection, care, and maintenance of archives, records, and manuscripts. B. Specific measures, individual and collective, undertaken for the repair, maintenance, restoration, or protection of documents.<sup>xiii</sup>

Ratcliffe Report, 1984

Strictly, all the steps taken to protect materials, that is including conservation and restoration, but often used in reference to the treatment of materials on first entering the library; it is preventive rather than remedial.<sup>xiv</sup>

## **Transition to Digital**

IFLA, 1986

Includes all managerial and financial considerations including storage and accommodation provisions, staffing levels, policies, techniques and methods involved in preserving library and archive materials and the information contained in them.<sup>xv</sup>

Feather, Matthews, and Eden, 1996

The managerial, financial and technical issues involved in preserving library materials in all formats--and/or their information content--so as to maximize life.<sup>xvi</sup>

## **Digital**

SAA, 1997

Preservation of digital information is not so much about protecting physical objects as about specifying the creation and maintenance of intangible electronic files whose intellectual integrity is their primary characteristic. Preservation goes beyond saving such media as optical disks or magnetic tape; the access system itself must be preserved.<sup>xvii</sup>

ICA, 1997

An electronic record is preserved if and only if it continues to exist in a form that allows it to be retrieved, and, once retrieved, provides reliable and authentic evidence of the activity which produced the record.<sup>xviii</sup>

Kenney and Rieger, 2000

Digital Preservation means retaining digital image collections in a usable and interpretable form for the long term. While “long-term” suggests an indefinite future, David Bearman interprets it more usefully as ‘retention for a period of continuing value.’<sup>xix</sup>

Archival and library definitions have shifted from the physical care and protection of materials to retaining them in retrievable form for an indefinite amount of time. In the paper-based information world, librarians and archivists sought to preserve books and documents for 500 years or more. As is apparent from both the study respondents, and the professional literature, professionals now think about maximizing “useful life” or preserving digital documents “forever” through emulation or forward migration, but without the emphasis on a specific number of years.

Further analysis of our data indicates that archivists and librarians view preservation through different lenses. This reflects a fundamental difference in the archival and library professions. Librarians tend to be custodians of printed materials that are not unique. Librarianship carries custodial responsibilities, but--with the exception of special collections--missing or damaged items can usually be replaced. Therefore, librarians often view their materials in terms of immediate utility. In the archival arena, when a record is gone, it is really gone and cannot be replaced--whether that is due to an accident or a disposal schedule. Archivists have responsibility for one-of-a-kind records, which are lodged in a repository. In current practice, the repository and the object cannot be divorced. This relationship differs from libraries and printed materials. In archives, long-term accessibility to the records may be mandated by legal warrant and business processes, and more broadly, by societal memory. The impact of electronic records may have an effect on the requirements that the repository and the object remain together in archives. In the digital environment *both* librarians and archivists have responsibility for documents and records that are born digital. These digital assets are susceptible to obsolescence and incompatibility.<sup>xx</sup> Therefore, the integrity and authenticity of digital objects is of mutual concern to both professions. As librarians and archivists work closely on long-term retention strategies, the definition of preservation may shift to accommodate both professional perspectives.

### *Costs*

In the section on Costs, we asked: “What do you estimate are the costs to preserve the records?” Responses included staff, equipment, space, energy and other related costs.

In essence, we were asking, what is it going to cost the institution to preserve, maintain, and provide access to electronic records? We thought this was an important question because for many institutions and projects, knowing what the bottom line is, is THE major factor which

influences decision-making, and determines goals and objectives, as well as the strategies to meet them. Knowing about costs helps institutions lobby with parent institutions or funding agencies. What is the role of costs in situations where because of legal requirements, archives do not have a choice about what they preserve or even how they preserve records? The majority of the managers we interviewed are gathering financial data now and plan to report costs as part of their projects' results. Only a few projects are far enough along to have developed cost figures. The interviewees ranged from large national archives, to projects developing testbeds. The costs for electronic record preservation ranged from \$10,000 to \$2.6 million per year. Cost categories include staff, consultants, facilities, equipment, storage system monitoring, staff access and research and development.

Most of the projects are currently funded through initial allocations, and some of these figures reflect the impact of early research and development costs, which could also account for the wide range of costs. In fact, as one respondent said, the costs for his project might be reduced by as much as half during the following year. This question will be followed up as part of the second phase of the research interviews. It will be interesting to see what the forecast figures for preservation, storage and staffing actually turn out to be, especially when the initial costs of research and development are reduced over time.

At the time of these interviews, none of the respondents had yet gathered enough information to determine the categories of preservation costs or cost modeling protocols.

Sources of funding include various government agencies, EU (European Union), NSF (National Science Foundation), NPACI (National Partnership for Advanced Computational Infrastructure), NEH (National Endowment for the Humanities), NHPRC (National Historical Publications and Records Commission), NARA (National Archives and Records Administration), and JISC (Joint Information System Committee). As always, the question remains as to what extent the source(s) of funding have shaped the research agenda and from there, the future.

The follow-up study will gather data on the further development of a preservation cost model. So far, cost modeling for digital projects has received scant attention. The present focus appears to be on budgeting for digital conversions rather than preserving authentic electronic records. In addition, there is scant literature in the area of cost models for born digital electronic records. Two exceptions are studies by Hendley and by Russell and Weinberger. Hendley, in his report on the *Comparison of Methods & Costs of Digital Preservation*<sup>xxi</sup>, provides a "Table of Digital Preservation Cost Elements" which was compiled by Neil Beagrie, Daniel Greenstein, and the Arts and Humanities Data Service. In it, the cost elements involved in developing and preserving digital collections are keyed to the life cycle stages of a digital resource.<sup>xxii</sup> In their article, *Cost Elements of Digital Preservation*, Kelly Russell and Ellis Weinberger posit that the ongoing costs of digital preservation span a more extended timeframe than traditional preservation and will therefore require resource commitments of a different nature. Different strategies may necessitate different costing time frames and schedules. Russell and Weinberger state that current cost models have yet to reflect this more complex environment. They further state that, "The creation of a digital object is the true starting point for digital preservation."<sup>xxiii</sup>

To estimate a budget for image acquisition, Anne Kenney and Oya Rieger refer to the “RLG Worksheet for Estimating Digital Reformatting Costs” in their book, *Moving Theory into Practice: Digital Imaging for Libraries and Archives*.<sup>xxiv</sup> The Worksheet, in combination with an assessment of costs derived by Cornell’s Department of Preservation, identified costs for image acquisition in six cost categories. These costs include personnel, equipment, cataloging, supplies, contingency and overhead/indirects.

To facilitate the development of a preservation cost model, a number of categories may be adapted from traditional cost models. These categories might include: costs for providing access to the materials; costs related to long-term creation and maintenance of digital materials, production of metadata, personnel, equipment, cataloging, supplies, contingency (e.g. emergency/unforeseen events), overhead, administration and, research and development.

One respondent provided information about plans to form a consortium of institutions to form a National Preservation Center. This idea should be explored not only because of its potential for cost-effectiveness of preservation, but also for the opportunities to enrich the library, archival and museum professions, which may occur as a result of providing a forum for communication across institutional settings and domains.

In a speech for directors of the Association of Research Libraries, Clifford Lynch stated,

The fundamentally hard things about managing bits into the future mostly aren’t technical; they’re economic and organizational. Bits need care and feeding. They don’t do well with benign neglect. This means that we need to come up with financial models to keep these bits cared for and healthy as they are migrated into the future. We don’t lose a lot of bits to technical failures in a well-managed environment, but we lose a lot due to financial or organizational failures to maintain that well-managed, caring environment on a continual basis.<sup>xxv</sup>

We include this quote to emphasize that technical processes cannot be separated from economic issues. The library and archival professions have not fully grappled with the economic influences on preservation decisions. It is necessary to identify concepts and approaches for evaluating the full economic impact of long-term retention. Institutional, national, and multi-national policies must be put in place to assure preservation in perpetuity.

## **Preservation Policy**

We concluded the survey with the following three questions about policy:

1. Do you have a general preservation policy that includes records in electronic form?
2. If not, do you have a policy for reformatting, refreshing, migrating, emulating, or bundling data to newer technological platforms?

3. Please describe any policies you might have that relate to preservation of electronic records.

Only three of the projects/institutions indicated having policies in place; two others are revising existing policies to include electronic records; and one is currently developing a policy that includes multiple media. Two of the research projects indicated that policy development would be an outcome of their research.

During Round 2 of the preservation survey, we will be interviewing key informants/experts who may shed more light on policy issues. We will try to ascertain whether or not international concern about the longevity of digital information is being followed up in policymaking arenas. We suspect that policy is lagging far behind the development of standards, because the development of good public policy requires the appropriate political climate as well as the cooperation of numerous stakeholders. Further, there must be a legal environment that enables the preservation of digital information. Yet laws may vary. For example, the Berne Convention and US copyright law have significant differences between them. These types of discrepancies may impede the development of consistent, rational public policy.

## Conclusion

**At present, the interviews indicate three broad themes. First, that the perception of what preservation is goes beyond library and archival practice to the media being preserved. Because electronic material is inherently ephemeral, and the timeframe involved to preserve and provide access to this material extends to perpetuity, we expect that traditional definitions of preservation may not apply. Indeed, a shift is already apparent.**

**Second, the rush to develop the technological processes necessary to preserve authentic electronic records appears to be at the expense of directly addressing cost and policy issues at the start of projects. One respondent, who is fully funded by his government, put it succinctly when he said, “We haven’t yet been asked to measure costs! We don’t need to justify costs. Fixed costs are unknown.” Another respondent said, “The result will be cost determinations.” And a third answered that costs “should be a result of the current testbed project [that they are engaged in].” We believe that the problems posed by preserving authentic electronic records permanently (or as long as possible), requires the development of a cost model, which will be unique and not a hybrid of existing digital conversion cost models. We agree with Hendley, Russell, and Weinberger that preservation begins at the creation of the electronic material. A cost model for preserving authentic electronic records will need to reflect this perspective, which differs from the traditional preservation point of view. Costs, however, cannot single-handedly be used to justify *not* preserving otherwise valuable records.**

**We found that staff and equipment costs are the most consistent hard figures available so far, and of course, those will vary over time, which will ultimately connect with developing forecasting strategies. Many of the projects are nascent, and we suspect that for them, answering the survey questions was essentially a theoretical exercise. As the institutions and projects progress, we expect to be able to gather hard data during rounds 2 and 3 of the survey. By the conclusion of round 3, we expect to have a substantial amount of information about institutions and projects that will have been active for at least 3 years. From this, we hope to develop a cost model for preserving authentic electronic records, which can be applied to archives and libraries, and perhaps to other communities of practice.**

Last, the lack of preservation policies in place is a distinct gap in the research design of many of the projects and possibly reflects a lack of commitment among the stakeholders in institutions. What is the reasoning behind developing policy as an end result of a project, instead of concurrently with its progress? We suspect that meeting the technological challenges of preserving electronic records is more of a priority within these institutions than developing policy and wonder whether, as a result, the overall progress in this new arena will be more uneven than is necessary. Several institutions that responded to our survey have had active programs for a long time and we note that often policy evolved, rather than being strategically planned. It is practically impossible to set policy 100% at the outset of a project—especially one in such a complex area as the preservation of electronic records. Policy will naturally evolve rather rapidly at the outset of a program when the practitioners encounter new, possibly unanticipated features of the program that require policy decisions. As the program matures, and even while it is still developing, policy will concomitantly need to be re-thought or newly conceived. In fact, policy must also drive technological development. When the program is in “full swing,” policy will have reached a point at which it is now well thought out, though still subject to modification, as the program requires.

In the subsequent phases of our survey, we hope to explore not only the “why” behind the positioning of policy development within the institution, but also the development of its content. We want to explore the role of the stakeholders and the influence of the legal and political environments which provide the context in which policy is formed.

We note that one project we included in round 1, has discontinued its research because funding ran out. This particular project was unique in that it was exploring the preservation of multimedia material. The gap in potential knowledge that could have been disseminated is a loss. But it is also a reminder that the adoption of any new technology depends on politics, funding, and timing.

Round 2 of the preservation survey will focus on expanding our knowledge in several areas. These include staffing and personnel – where are future specialists in preserving electronic records going to come from? Another expanded area in the survey is cost activities. Because of the nascence of some aspects of the programs we studied, such as charges, access, reproduction and copyright, we were able to gather very little substantive information. As a result, these sections will probably drop out of round 2. However, we will re-evaluate their inclusion for

round 3. As well, we intend to explore in more depth, why certain questions did not apply to some respondents.

No matter which preservation method is chosen, cost will become a factor in making a management decision regarding preservation of electronic records. We have also expanded the area of our survey that asks for information on description/documentation of preservation processes, as well as the section on preservation policy. In addition, we have developed a second survey instrument that we will use to interview key informants/experts, whom we define as “individuals who provide useful insights in the fields of preservation and/or archives.” These experts may work in a variety of settings including – but not limited to – universities, government archives, foundations/granting agencies, industry, professional organizations, or think tanks, or who serve as consultants. Most have extensive national and international experience that enables them to provide the long view of preservation as well as placing them at the forefront of their professions.

\* \* \* \* \*

We began our study by asking four research questions:

1. 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?
2. What is the meaning of *preservation*?
  - a. Does the meaning change when it is applied to electronic rather than paper-based records?
3. Will current strategies for preserving electronic records ensure longevity and authenticity?
4. How are costs for the preservation of electronic records derived? Have effective cost models been developed?

Our survey provided us with only partial answers to Questions 1 and 3. We identified a number of preservation techniques that are currently in use—including migration, emulation, and robotics—but not one of these techniques could be considered to meet the conceptual requirements for authenticity. Until these methods are further developed and standardized, they cannot be relied upon to ensure the long-term preservation of electronic records. As for question 1b, “Which methods of long-term preservation need to be developed?,” they all do. It is still too early in the development of all these techniques to fully evaluate them. Regarding question 1c, “Which of these methods are required or subject to standards, regulations, and guidelines in specific industry or institutional settings?,” there is no simple answer. The projects represented in our survey are developing standards and guidelines. Some of the institutions we interviewed

are waiting to see the results of these projects before committing to a particular strategy. We hope to be able to answer question 1 more fully in subsequent rounds of this research.

Question 2 yielded richer results. It is clear that professionals are revising their definitions of preservation from a once-and-forever approach for paper-based materials to an all-the-time approach for digital materials. Preservation must now accommodate both media and access systems. Finally, while we once tended to think about preserving materials for a particular period of time—for example, permanent/durable paper was expected to last for five hundred years—we now think about retaining digital media for a period of continuing value.

Meaningful answers to Question 4, regarding costs for the preservation of electronic records, must also wait until rounds 2 and 3. Our survey revealed that in the rush to develop the technological processes necessary to preserve authentic electronic records, cost issues have often been pushed aside. This is in part because ample government and foundation funding is allowing some institutions to defer cost modeling. Many respondents reported that they are beginning to study the cost implications, and we hope to gather more information in the next round.

As a result of the information we will gather over the next two years about evolving preservation practices, we expect to strengthen the foundation underlying the development of the preservation function model, particularly those aspects which concern preservation, storage, and access to authentic electronic records over time. We also hope to provide insights which will contextualize the work of projects and institutions around the world, and which will ultimately provide a pool of knowledge that will benefit us all.

\* \* \* \* \*

When John Steinbeck completed *Of Mice and Men*, he described it to his publisher as an experiment, adding, “don’t publish it if you don’t like it.” So unsure was the author of his work that he did not even want to read the proofs. Christopher Morley, describing for *Book of the Month Club News* the publication of Steinbeck’s book, wrote that “[I]n just such casual ways, in this our world of obliquity and squint, do masterpieces happen.”<sup>xxvi</sup> “Obliquity and squint” beautifully captures the notion of looking at something without fully understanding or seeing it. It is as apt a description of electronic preservation as it is a description of the chance publication of an enduring work of literature. In this period of incunable electronic information, it is difficult to understand all the potential or all of the pitfalls of our newest cultural heritage. Until we do, however, electronic preservation itself will be seen as oblique.

---

<sup>i</sup> Howard Besser, "Digital Longevity," in *Handbook for Digital Projects: A Management Tool for Preservation and Access*. (Andover, MA: Northeast Document Conservation Center, 2000): 156.

<sup>ii</sup> See, for example, *Research Libraries Group News* 52 (Spring 2001): 3, 8-9.

For various definitions of *authenticity* and *authenticity of digital information*, see Nancy Brodie, "Authenticity, Preservation and Access in Digital Collections," *The New Review of Academic Librarianship* 6 (2000):225-238.

Anne Gilliland-Swetland, *Enduring Paradigm, New Opportunities: The Value of the Archival Perspective in the Digital Environment*. (Washington, D.C.: Council on Library and Information Resources, February 2000): Gilliland-Swetland, 11-12.

Gilliland-Swetland, 13.

Margaret Hedstrom and Sheon Montgomery, *Digital Preservation Needs and Requirements in RLG Member Institutions*. (Mountain View, CA: Research Libraries Group, 1998).  
<http://www.rlg.org/preserv/digpres.html>

Available: In Anne R. Kenney and Oya Y. Rieger's *Moving Theory into Practice: Digital Imaging for Libraries and Archives* (Mountain View, CA: Research Libraries Group, 2000), the contributors cover particular topics (e.g. "What Users Want from Digital Image Collections," "Benchmarking for Conversion," "What About Copyright?"), but the book doesn't compare practices across institutions.

Reagan Moore, et al., "Collection-Based Persistent Digital Archives – Part 2," *D-Lib Magazine* 6.4 (April 2000): 1.

See also, Stewart Granger, "Emulation as a Digital Preservation Strategy," *D-Lib Magazine* 6.10 (October 2000): 1-12 We were not able to establish contacts at two of the sites (the National Library of Sweden and the Department of Computer Science at Stanford University).

Ross, Seamus and Ann Gow. *Digital Archeology: Rescuing Neglected and Damaged Data Resources, A K/JISC/NPO Study*. (Glasgow: Humanities Advanced Technology and Information Institute (HATII), University of Glasgow, February 1999): iv. <http://www.hatii.arts.gla.ac.uk>

Hedstrom and Montgomery, 5.

These have been slightly copyedited.

Frank Evans, et al. "A Basic Glossary for Archivists, Manuscript Curators, and Records Managers." *The American Archivist* 37.3(July 1974): 415-433.

F.W. Ratcliffe with D. Patterson. *Preservation Policies and Conservation in British Libraries: Report of the Cambridge University Library Conservation Project*. Library and Information Research Report, 25. (London: The British Library, 1984).

J.M. Dureau and D.W.G. Clements, *Principles for the Preservation and Conservation of Library Materials*. IFLA Professional Reports 8. (The Hague: IFLA, 1986).

A definition that Feather, Matthews, and Eden came up with based on broadly accepted definitions. See John Feather, et al. *Preservation Management: Policies and Practices in British Libraries*. (Aldershot, England: Gower Publishing, 1996), p. 5.

---

<sup>17</sup> *Statement on the Preservation of Digitized Reproductions*. (Chicago: Society of American Archivists, 1997).

<sup>18</sup> International Council on Archives, *Guide for Managing Electronic Records From an Archival Perspective*. ICA 8 (Paris: ICA, February 1997): 32.

<sup>19</sup> Anne R. Kenney and Oya Y. Rieger, *Moving Theory into Practice: Digital Imaging for Libraries and Archives*. (Mountain View, CA: Research Libraries Group, 2000), p. 135.

<sup>20</sup> Jeff Rothenberg is now advocating use of the term “inherently digital” to characterize executable, dynamic, and interactive materials. See his *Using Emulation to Preserve Digital Documents*. The Hague: Koninklijke Bibliotheek, July 2000.

<sup>21</sup> Hendley, Tony. *Comparison of Methods & Costs of Digital Preservation*. (1998) p. 96.  
<http://www.ukoln.ac.uk/services/elib/papers/tavistock/hendley/hendley.html> (downloaded October 2000)

<sup>22</sup> Hendley, p. 65.

<sup>23</sup> Russell, Kelly and Ellis Weinberger. *Cost Elements of Digital Preservation*.  
<http://www.leeds.ac.uk/cedars/documents/CIW01r.html> [downloaded August 2000].

Kenney and Rieger, p. 166.

<sup>25</sup> Lynch, Clifford. “Strategic Issues: Technology, Trends and Solutions.” In *Preserving Digital Information*, Vantage Point series, pp. 3-4. EBSCO Subscription Services, 2000.

<sup>26</sup> Christopher Morley as quoted in, Al Silverman, ed., *The Book of the Month: Sixty Years of Books in American Life*. (Boston: Little, Brown, 1986): 49.