

**DIGITAL PRESERVATION ARCHIVING WORKSHOPS: FOCUSING THE AGENDA****PROPOSAL SUMMARY:**

The proposal requests funding for five two-day workshops on specific problem topics of digital archiving, with participants both invited and self-selected through competition. Each workshop will devote itself to the particular topic with the aim of noting what progress has been made so far, specifying what the major problems in the topic are, and what would be fruitful means of investigating the problems and solving them. Reporters will produce summary papers outlining conclusions for both print and World Wide Web publication in prompt, convenient formats. The aim is to produce research and action agendas to guide already-active digital library creators in furthering effective digital preservation and archiving. The Sponsors for the Project and the workshops will be the Association of Research Libraries, the Coalition for Networked Information, and the Commission on Library and Information Resources.

Participants will include knowledgeable individual practitioners from organizations active either in planning or in creation of digital repositories. Participants will also include self- (or other-) nominated individuals. Workshop attendance should comprise approximately 30-50 people. This will be large enough to assure multiplicity of views and ability to work in smaller breakout groups, but not so large as to prevent considerable informal exchange. Participants should expect to fund their own expenses in attending the workshop (thus this is a matching fund proposal with respect to most participants). The Project, using grant funds, will fund a small number who could not otherwise come.

At each workshop the goals will be:

- To identify what research or development on the topic has been done, if any, and where clear consensus may already have been reached.
- To find a consensus on the most pressing research and action agendas for potential practitioners;
- Identify and prioritize, if possible, the next action steps that should be taken in the topic area.
- Where possible, identify which organizations or institutions are the best candidates for proceeding, including pressing for volunteer organizations at each workshop.

The workshops will each last for approximately two days (probably requiring two overnight stays for most attendees). The workshops will include invited speakers, breakout sessions and plenary sessions.

Summary statements reflecting each workshop outcome will be prepared immediately after each workshop. These statements and documents will be made available on the network through a Web site maintained by either the Director or one of the Advisory Committee participants; the editing and presentation values will be at minimal levels to expedite their appearance. At the conclusion of the Project a more formal version will be published both in print form and on the Web and made widely available.

**TABLE OF CONTENTS**



**DIGITAL PRESERVATION ARCHIVING WORKSHOPS: FOCUSING THE AGENDA**

NEH Cover Sheet..... i

PROPOSAL SUMMARY..... 1

TABLE OF CONTENTS..... 2

BACKGROUND..... 3

METHODOLOGY..... 5

    1. Planning

    2. Workshops

    3. Products

PROPOSED WORKSHOPS..... 6

    1. Archival contents and redundancy

    2. Technological Obsolescence and Migration techniques

    3. Current and immediately future hardware technologies

    4. Authenticity (integrity; intellectual preservation)

    5. Repository certification

    6. Intellectual Property

PARTICIPANTS..... 8

PROJECT STAFFING..... 9

VENUES..... 10

FUNDING AND BUDGETARY ASSUMPTIONS: ..... 10

NEH BUDGET SHEETS follow page 11..... 12-15

**APPENDICES**

I. BIBLIOGRAPHY.....16

III. PROJECT DIRECTOR VITA.....21

III. ACRONYMS.....23

IV. SPONSORING ORGANIZATION DESCRIPTIONS (ARL, CNI, CLIR)..24-29

V. LETTERS OF SUPPORT.....30-32

NEH Sheets

History of Grants.....33

[Original copy only:]

List of Project Participants  
    and Members of Advisory Board.....34

List of Suggested Evaluators.....35

**DIGITAL PRESERVATION ARCHIVING WORKSHOPS: FOCUSING THE AGENDA****BACKGROUND:**

The provision of information has long been the traditional responsibility of libraries, and the maintenance of information for the long term has been the distinguishing feature of research libraries. In the past several years as the electronic revolution has reached full force there have been many national and international projects recognizing the need to establish digital libraries (see bibliography). Outside the library and archival communities there has been almost no recognition of the importance of long-term preservation (here called archiving) of electronic information. In part this is because of a lack of library perspective, and in part because of the inherent difficulty of the topic. Within the library community there are several important digital library projects addressing archiving as a critical need, but they are uncoordinated and to date publicly almost mute.

The current proposal joins in recognizing the critical importance of archiving as research libraries take on the responsibility of organizing and providing electronic information. This proposal intends to aid the already active participants in working out what are the critical issues in digital archiving, and how those critical issues might best be attacked.

There is a need for a coordinated, international planning effort to further digital archiving. Good work has already been done in working toward the reality of digital libraries. Outstanding examples include:

- The RLG/CPA report on digital archiving of 1996;
- A number of the papers produced by the Commission on Preservation and Access over the past few years;
- The focus of the National Digital Library Federation on archiving as a critical digital library issue;
- The Research Libraries Group ARCHES project, arising out of its "Studies in Scarlet" project, to establish a model for a digital library;
- The recent OCLC announcement that it will archive electronic journals to which it provides subscription access;
- The eLib project in the United Kingdom, which with substantial government funding initiated by the Follett Commission is investigating digital library issues;
- The Australian PADI (Preservation of Access to Digital Information) Working Party, formed last year;
- The archival community's early focus on electronic preservation, carried on in part by the electronic records community's listserv on the topic;
- The six NSF-funded Digital Library projects, which are developing important bodies of knowledge with respect to large-scale data bases, access, interoperability, and image management (though they have as yet only glancingly addressed archiving issues).
- Significant progress in increasing accessibility of networked information through development of the Dublin Core and the Handles initiatives along with IETF work on Uniform Resource Characteristics.

In spite of these important beginnings, no USA organization is yet pressing

the archival issues with an active, energizing, collegial plan intended both to create a significant archiving capability and to include a broad spectrum of the research library community as the players they will inevitably be. In fact there is likely to be significant agreement as to the strategic goals of archiving, including such matters as medium preservation, protection against technological obsolescence, and assurance of information integrity. But for a variety of reasons, including wholesome competition, the difficulty of managing large-scale projects with many players, the perceived scarcity of intellectual resources, and the very real scarcity of financial resources, most of the US projects have proceeded behind the scenes and with little public discussion of their efforts. Without coordination, this healthy multiplicity of activity (it would not be fair to call it redundant) could follow divergent and therefore counterproductive paths.

Communication between the major players is essential to develop consistent programs and standards, as well as best practices that have the consensus of the best participants. In this way the using community (certainly including scholars but many others as well) can have clear expectations about the nature of the information they examine when they approach the digital libraries of the future.

A great deal will be learned as these multiple project participants communicate about what has worked and what has presented difficulties. In addition, there is great value in the library projects collaborating with practitioners in other fields, for example the social science data archivists at ICPSR in the United States, the electronic records practitioners within the archivist community, the Data Archive at the University of Essex, and the computer scientists working with large-scale data bases in the NSF projects; and with developers in other countries, most notably in the U.K. and Australia.

Perhaps most important is that the research and development agenda is not yet clear. Many informed and knowledgeable librarians, and others, are speaking of digital libraries and of electronic archiving. But there is not yet consensus on what are the strategic archiving goals to be achieved, nor on the tactics of attacking the problems that stand in the way of those goals. The nature of the problems in fact remains obscure; certainly there is no settled view as to what they are.

Establishing the R&D agenda is the goal of the proposed Project. Granting agencies and major national (and international) institutions are prepared to fund projects to further digital archiving, but they need to know they are applying their funds toward the most useful outcomes. Practical research in digital archiving is needed to determine best practices, technology tools, effects of techniques, and the like. But even before economical, focused research can begin there needs to be some consensus as to what are the most significant problem areas and what are the hoped-for goals.

Such a consensual effort should not proceed at the relatively slow traditional pace that has characterized the methodical, collegial development of high-quality library standards. This effort must be complemented by the action

orientation of the computing community. The development of digital archives should be from the user orientation of libraries, but also from the IETF motto of prompt implementation on the basis of "rough consensus and working code."

This Project recognizes that the research agenda is not yet clear and that there is a need to clarify it. The Project proposes that conversations between interested parties will help to do this, and thus proposes to effect such conversations in a timely fashion that will match the national and international need.

**METHODOLOGY:**

The NEH calendar calls for support announcements not to be made until April, 1998. The Council on Library and Information Resources (CLIR), one of the Project's sponsoring organizations, has committed itself to providing an initial grant to allow planning and first steps to begin (see appendix).

**1. Planning:** With the aid of the planning grant from CLIR, the Project will take the following steps:

- Organize the Advisory Committee, comprising senior representatives from the Sponsors (described below): CLIR, the Association of Research Libraries (ARL), and the Coalition for Networked Information (CNI). Initial activities will include holding a first meeting, establishing electronic communication mechanisms, establishing budgetary and programmatic oversight mechanisms, laying out a calendar, and the like.
- Refine or even revise the workshop topics in this proposal, though major changes are likely to be few if any.
- Publicize the likelihood of workshops as described in this Proposal so as to increase awareness of the possibility of common action.
- Through contact with many of the leading individuals in existing archiving projects, the Project Director will begin developing a list of potential workshop invitees (both organizations and individuals), casting the initial net as broadly as possible. Contacting those persons, he will expand the list further. The Advisory Committee with the Project Director will determine the final list of invitees.
- For each of the major workshop topics, the Project Director will begin developing a statement of what is known and has been done so far and by whom. This will not be comprehensive but instead compiled with the intent of finding an appropriate person to provide a fuller summary at the outset of each workshop as well as an advance paper to jump-start the workshop discussions.

**2. Workshops:** As funding becomes available, organize five topical workshops on digital archiving as described below. At each workshop the goals will be:

- To identify what research or development on the topic has been done, if any, and where clear consensus may already have been reached.

- To find a consensus on the most pressing research and action agendas for potential practitioners;
- Identify and prioritize, if possible, the next action steps that should be taken in the topic area.
- Where possible, identify which organizations or institutions are the best candidates for proceeding, including pressing for volunteer organizations at each workshop.

The first step after funding is granted will be to publicize the Project and the likely calendar to allow potential attendees to begin planning.

The workshops will each last for approximately two days (probably requiring two overnight stays for most attendees). The workshops will include invited speakers, breakout sessions and plenary sessions. Speakers and other active participants will be determined by the Advisory Committee and Project Director working closely with leaders in the subject area of each topic.

**3. Products:** Summary statements reflecting each workshop outcome will be prepared by the Project Director (or a workshop volunteer) immediately after each workshop. More discursive statements may be developed as agreed upon by participants at each workshop (part of the agreement will be to provide such statements with more than deliberate speed).

These statements and documents will be made available on the network through a Web site maintained by either the Director or one of the Advisory Committee participants; the editing and presentation values will be at minimal levels to expedite their appearance. At the conclusion of the Project a more formal version will be published both in print form and on the Web, again under the auspices of one or more of the Advisory Committee members, and made widely available.

#### **PROPOSED WORKSHOPS:**

Following are proposed workshop topics, with related questions for which further direction is needed. The five workshops will be selected from among the six described below, after wide consultation. Each workshop should include a review of present activities in the topic area, as described above. The planning for each workshop will no doubt result in further questions and subtopics:

#### **1. Archival contents and redundancy ("collection development"): What is the necessary agenda?**

What needs to be archived? What doesn't? What distinguishes what research libraries archive from what other organizations archive (e.g. ICPSR, NARA)?

To what extent are existing collection development practices valid in the archiving environment? What needs to change?

What do libraries need from parent institutions in this regard? What is possible? What role will universities play? Will universities agree to serve as archives?

What are the repository redundancy needs? Issues include storage costs and network bandwidth, disaster issues, geographic concerns, loci of authoritative sites, etc.

What are repository agreements likely to be based on? Consortial arrangements, political convenience (e.g. statewide arrangements), common vendor relations, subject area agreements, etc. What are the pros and cons of these possibilities?

Does a survey of present activities in this area make sense, and if so, how should it be structured and presented?

## **2. Technological Obsolescence and Migration techniques: What is the necessary agenda?**

Just in case or just in time? (Migrate all electronic data through all technologies or wait for demand to determine what is migrated?)

What are practical technology emulation possibilities?

Migrate the information object only or migrate its tool for use? Link them together or provide only implicit linking information?

What are the major issues in migration?

## **3. Current and immediately future hardware technologies: What is the necessary agenda?**

What are effective current storage media?

To what extent should future refreshing/migration be taken into account in choosing current storage media?

Is vendor cooperation possible?

What are the prospects for future hardware technologies as archiving media (HD-Rom, ceramics, etc.)?

What is the interplay between hardware permanence and software permanence? What are the strategic consequences?

Online, nearline and offline: when to use what.

## **4. Authenticity (integrity; intellectual preservation): What is the necessary agenda?**

What are potential tools and/or mechanisms for assuring integrity? What is the experience so far? What are the user burdens? How well do they work? What infrastructure do they require? What costs?

What needs to be assured, and what need not be? What principles can be established in this area so that constant individual decisions are not necessary?

What does integrity mean? Can content assurance be provided though presentation modes change? Does that matter?

What does integrity mean in a migrating environment? What is the fundamental "content" to be preserved? What can afford to be lost?

Are there differences in document type that affect the integrity mechanisms?

What are the relations between integrity mechanisms and version control? Do they matter?

## **5. Repository certification (NOTE: this workshop best held after the others): What is the necessary agenda?**

What should be the standards for certification as a repository?

Should institutions self-certify? If not, who should, or how should they be?

Should products be part of the certification process, or process?

Is research library repository certification a stand-alone issue, or part of a larger issue (are research libraries one component of an archiving repository community)? Should certification standards be different for different communities?

**6. Intellectual Property: What is the necessary agenda?**

What are the intellectual property constraints placed on preservation distinct to the preservation activity (i.e. other than the constraints and requirements of present day to day use)?

What are the potentials for libraries gaining some form of intellectual trusteeship over information, whether in partnership with producers or in the event of their failure to act ("fail-safe" triggers)?

What intellectual property legislation, or international agreements (e.g. through WIPO), or commerce conventions, are needed to facilitate preservation of the human record?

How can intellectual property markers be related to the preserved information (i.e. via metadata such as PICS, Warwick Framework, or Handles implementations)?

**PARTICIPANTS:**

Participants will be of at least two kinds:

1. Participants should include knowledgeable individual practitioners from organizations active either in planning or in creation of digital repositories. Obvious potential organizations are RLG, OCLC, NDLF, Library of Congress, NINCH, CNRI, representatives of the NSF-funded digital library projects, National Library of Australia, National Library of Canada, UKOLN (and related, e.g. British Library R&D), the Digital Libraries Conference Planning Group. Others, depending upon the topic, could well include ARL, scholarly societies, AAUP presses and other publishers, data base owners, legal experts, senior university representatives, government officials and technology experts.
2. Participants should include self- (or other-) nominated individuals. In this early stage of development there is likely to be a number of people who have contributions to make though they may not be at institutions actively engaged in projects, or may not themselves be part of the project. Potential participants could include knowledgeable librarians, library school faculty, researchers and faculty in computing and information disciplines, academics with special concerns for the longevity of their knowledge base, representatives of commercial solutions, publisher representatives, and the like.

A call for participants should be published sufficiently in advance to allow for applications and their timely review. Applications should include vitae, brief justification for attendance, constituencies explicitly or implicitly represented, and proposals for disseminating information about the workshop discussions. Review of participant proposals should be by the Advisory Committee aided by a small number of reviewers (varying for each workshop). This committee should do its work electronically.

Workshop attendance should comprise approximately 30-50 people. This will be large enough to assure multiplicity of views and ability to work in smaller breakout groups, but not so large as to prevent considerable informal exchange

(BOFs, etc.). Determination of size for each workshop will be by the Advisory Committee.

Participants should expect to fund their own expenses in attending the workshop (thus this is a matching fund proposal with respect to most participants). The Project, using grant funds, will fund a small number who could not otherwise come. There is no necessary relation between self-nomination or invitation, on the one hand, and Project funding of attendance costs, on the other.

All participants should commit in advance to producing written outcomes of workshops in publishable form promptly after attendance; this assignment may end up being delegated to a relatively few people, but those who are chosen should be under commitment to produce documentation of the event.

A few observers may be invited from organizations whose mission assumes an overview of such proceedings, e.g. CLIR, NEH, ARL, CNI, NSF and the like, to attend also at their own expense.

#### **PROJECT STAFFING:**

The sponsoring organizations (Sponsors) will be:

- The Association of Research Libraries (ARL);
- The Coalition for Networked Information (CNI);
- The Council on Library and Information Resources (CLIR).

Though the Sponsors do not have operational programs, each is a leader in the development of thinking intended to result in effective digital research libraries. The Sponsors will form an Advisory Committee to the Project which will comprise senior executive officers or member officials, as each organization decides is appropriate. The Project should not be seen as a single-person initiative or as a peripheral venture. The Advisory Committee will provide oversight and direction to the Project Director (PD), both programmatically and organizationally. The critical decisions on topics, attendees, publicity and the like will be made in a collegial fashion among the three organizations and the Project Director. The PD will provide regular reports to the Advisory Committee, including financial and planning reports before and after each workshop.

The Project Director will be Peter S. Graham, Associate University Librarian for Technical and Networked Information Services at Rutgers University Libraries. Mr. Graham has written and spoken on issues of digital preservation for several years, and recently won an award for an article on digital research libraries. He has been active in CNI and RLG planning activities, and more recently in ARL digital preservation program planning, and has participated in Dublin Core metadata activity. He has technical computing background as well as library experience. He is personally familiar with many individuals now involved in digital library activities. He has been a manager of several significant operations at large institutions and will be able properly to conduct the activities of the Project (see Appendix for his vita). The Rutgers University Libraries have approved his work on this Project and the University is prepared to provide its base of operations.

The Project Director will be responsible for working with the Advisory Committee and other major players with an interest in digital preservation and archiving matters. He will assure that workshop agendas are appropriate and well prepared, that proper publicity is provided, that actual workshop organization is effectively prepared and implemented, that summaries are developed and written, and that proper publication ensues. The Project Director will be responsible for the organization and accomplishment of each workshop. Decisions on speakers, workshop segment leaders, venues, and other substantive and organizational arrangements will be jointly the responsibility of the Project Director and the Advisory Committee.

Staff assistance will be needed at the Project Director's home institution; this will be provided through a small amount of clerical assistance and through the assistance of a library school student working as a graduate assistant (from the Rutgers School of Communications, Information and Library Science). The assistants will work on physical arrangements of workshops and of publications, preparation of materials for attendees and for workshops, advertisements and calls for participation (mostly electronic), project Web page maintenance, etc. The Project will benefit by working closely with staff at national organizations accustomed to organizing similar activities, e.g. the Sponsors themselves. Staff assistance at the workshops will be needed to assure smooth functioning; probably from the sponsoring organizations. Their travel expenses (but no other remuneration) are shown in the budget.

**VENUES:**

The workshops will be held in varied locations within the USA convenient to transportation and to loci of major organizations active in the topic area, e.g. hotels or university environments in San Francisco, Washington DC, Chicago, Rutgers/New Brunswick (near Newark Airport), New York City, etc. Where meetings are held in hotels (the most likely locations), the necessary meeting spaces may be partially covered by paid rooms of attendees, but as several rooms are necessary for breakout sessions there are likely to be further costs.

**FUNDING AND BUDGETARY ASSUMPTIONS:**

The following budgetary assumptions are provided to assist in interpretation of the formal budget sheets attached:

- There will be five workshops.
- There will be approximately fifty participants at each workshop, plus three international participants, the Project Director and two staff assistants.
- The attendance costs of the three international participants, the Project Director and two staff assistants will be supported by the Project.
- Ten of the domestic participants at each workshop are to be funded by the Project, at approximately \$1,000 each (airfare, hotel and per diem); all other participants are to find their own attendance costs; transportation costs are assumed to approximate \$600, hotel and per diem approximately \$400.
- Three international participants at each workshop are to be funded by the Project, at approximately \$2,500 each (airfare at about \$1,800, hotel and per diem about \$700; the total is based upon a one-week stay to take

advantage of cheaper air fares even though per diem costs might then be somewhat higher).

- Joint meal and coffee-break expenses will not be funded by NEH but will be provided through other funding to be arranged by the host institution.
- Meeting rooms will be partly provided by hotels or other meeting places based upon paid room attendance, but several breakout rooms and a plenary room will be necessary; the Project budgets for the cost of these rooms.
- Staffing at the actual workshops will be provided through staff volunteered by the Advisory Committee members, with expenses (not salary) paid by the Project.
- The Project Director will be aided in the logistics of planning by a graduate library school student and by a small amount of secretarial assistance at the home institution.
- Participants' own funding of their travel expenses is shown as cost-sharing in the attached budget sheets.
- Rutgers indirect cost recovery is at the 26% level reflecting that more than half of the Project will be conducted off-campus.
- In the budget sheets, Indirect Cost Recovery (IDR) amounts are shown for participants' own travel expenses, as required by the Rutgers Office of Research and Sponsored Programs. These IDR amounts are notional and do not reflect actual payments or budget transfers or any intentions in that respect.
- Outcomes publications:
  - Five intermediate productions reflecting each workshop will be produced.
  - The intermediate publications will be placed on the World Wide Web.
  - The intermediate publications and the final publication will require editing for style, consistency and expression, which is shown in the Project budget as requiring a consultant.
  - Web page publication will require additional editing and preparation by an experienced Web editor, which is shown in the Project budget as requiring a consultant.
  - The final publication will be in print form as well as electronic form; budget estimates for its production and wide distribution are estimates based on past experience.



## APPENDIX I: BIBLIOGRAPHY

[Excerpts -- Section D only -- from]

**Bibliography on Electronic Library Issues**<http://aultnis.rutgers.edu/texts/ElectLibBib.html>

June 23, 1997

[Excerpted] Contents of this page:



- A. Bibliography of Bibliographies on (and Links to) Electronic Library Issues.
- B. Syllabi and Course Reading Lists.
- C. Bibliography on Metadata Issues (including the Dublin Core), URx, and PURL.
- **D. Bibliography on Preservation and Archiving Issues.**
- E. Bibliography on Other Electronic Library Issues.

[...]

**D. Bibliography on Preservation and Archiving Issues**

1. Arches -- [RLG] Archival Server and Test Bed, a page describing the Research Libraries Group project at the basis of its digital library preparation. "The Arches project involves both an online repository for digital resources of all types and the software environment that makes this repository internationally accessible and responsibly maintainable. " (6/97)
2. ARCHIVISTS: a bulletin board, or journal (it's not quite clear which) is edited by Brewster Kahle, the WAIS developer. The first issue (to 421 subscribers) appeared on the list on Aug. 13, 1996. It is concerned with electronic archiving matters, in particular the Internet Archive that Kahle's group says it is maintaining. Subscription to ARCHIVISTS may be effected by emailing to archivists-request@archive.org and providing as the message the following:  
subscribe archivists (optional email address).  
The mission of the Internet Archive is gathering, storing, and providing access to public materials on the Internet such as the World Wide Web, Netnews, and downloadable software. The collection, reaching ten terabytes, will provide historians, researchers, scholars, and others access to this vast collection of data, and ensure the longevity of the information.
3. Center for Electronic Records Home Page, a product of NARA (the National Archives and Records Administration). Included here as an example of current activity in a major area (9/3/96).
4. The Data Archive of the University of Essex.  
"The Data Archive at the University of Essex houses the largest collection of accessible computer-readable data in the social

sciences and humanities in the United Kingdom. It is a national resource centre, disseminating data throughout the United Kingdom and, by arrangement with other national archives, internationally. Founded in 1967, it now houses approximately seven thousand datasets of interest to researchers in all sectors and from many different disciplines."

Like ICPSR (q.v.), they have been working on preservation and archiving techniques of potential use for digital libraries. Denise Lievesley is the director. (6/97)

5. Day, Michael. "Extending metadata for digital preservation." See in C above.
6. Digital Library Resources and Projects, "a Library of Congress Internet Resource Page" in collaboration with the National Digital Library Federation. A list of activities in digital library construction, including NSF projects and including a partial list of conferences on the topic. (6/97)
7. Graham, Peter S., "Intellectual Preservation and the Electronic Environment," *After the Electronic Revolution, . . . : Proceedings of the 1992 Association for Library Collections and Technical Services President's Program*, ed. Arnold Hirshon (Chicago: ALA, 1993), p. 18-38.
8. Graham, Peter S., Intellectual Preservation: Electronic Preservation of the Third Kind (Washington, DC: Commission on Preservation and Access, March, 1994), 8 pp.
9. Green, Ann Gerken, and Joann Dionne, Preserving the Whole: A two-track Approach to Rescuing Data and Metadata (Dec. 20, 1996). This is a Yale University Library project to migrate machine-readable data (survey research data) along with its associated descriptors (codebooks), or metadata. It is one of the few studies of actual migration I have seen, and while brief (about 7 pages) is very informative (<URL:<http://www-cpa.stanford.edu/cpa/misc/preswhol.html>>). (4/22/97)
10. Haber, Stuart, and W. Scott Stornetta, "How to Time-stamp a Digital Document," *Journal of Cryptology* (1991) 3:99-111; also, under the same title, as DIMACS Technical Report 90-80 ([Morristown,] New Jersey: December, 1990). DIMACS is the Center for Discrete Mathematics and Theoretical Computer Science, "a cooperative project of Rutgers University, Princeton University, AT&T Bell Laboratories and Bellcore." The authors are Bellcore employees.

A brief account is in Barry Cipra, "Electronic Time-Stamping: The Notary Public Goes Digital", *Science* Vol. 261 (July 9, 1993), p. 162-163.

See also Dave Bayer, Stuart Haber and W. Scott Stornetta, "Improving the Efficiency and Reliability of Digital Time-stamping," *Sequences '91: Methods in Communication, Security, and Computer Science*, ed. R. M. Capocelli (New York: Springer-Verlag, 1993), p. 329-334.

11. Hedstrom, Margaret, "Preserving the Intellectual Record: A View from the Archives," Paper prepared for an International Conference on Networking and the Future of Libraries, UKOLN, Office for Library & Information Networking, Univ. of Bath, England, 19-21 April 1995 (typescript: 26 pp., references). [Contact author--NY State Archives--for copies. --pg]
12. ICPSR Home Page, for the Inter-university Consortium for Political and Social Research (ICPSR), located within the Institute for Social Research at the University of Michigan. Richard Rockwell, Director. ICPSR has been engaged for some time in the problems of preserving digital information over time. (6/97)
13. "Implications of Electronic Formats for Preservation Administrators," Newsletter Insert: *Newsletter*, Commission on Preservation and Access (Washington, DC: Nov.-Dec. 1993, No. 62), p. 1-2.
14. Internet Archive: see ARCHIVIST, above.
15. Jack Kessler's CARL bibliography of October 20-21, 1995. 'Annotated "Bibliography": *Preservation in the Digital Age*: A handout for "Preservation in the Digital Age: Whether To, What To, How To", A Session of the Third Annual Conference of CARL, California Academic and Research Libraries, San Francisco, October 20 - 21, 1995.'
16. Lesk, Michael, *Preservation of New Technology: A Report of the Technology Assessment Advisory Committee to the Commission on Preservation and Access* (Washington, DC: CPA, 1992; available from the Commission at \$5: 1400 16th S. NW, Suite 740, Washington, DC 20036-2217).
17. *Long Term Preservation of Electronic Materials*, British Library R&D Report 6238. A report of "a JISC/British Library workshop as part of the Electronic Libraries Programme organized by UKOLN 27th-28th November 1995 at the University of Warwick." Available in print and also on the Web at <http://www.ukoln.ac.uk/services/elib/papers/other/preservation/>. Papers by Peter Graham (on preserving the digital library), Denise Lievesley (on managing digital archives), Daniel Greenstein, Sir Anthony Kenny (on legal deposit), Margaret Hedstrom (on preserving digital information), Mirjam Foot (on preservation policies), Tony Hendley (on practical implications). (6/97)

18. Lynch, Clifford, *Accessibility and Integrity of Networked Information Collections* (Office of Technology Assessment, Congress of the United States, July 5, 1993; 107 pp.).
19. Lynch, Clifford, *The Integrity of Digital Information: Mechanics and Definitional Issue*, *Journal of the American Society for Information Science*, 45(10):737-744, 1994. This is a very useful article on issues revolving around authentication, integrity, intellectual preservation, and the like.
20. Mohlhenrich, Janice, ed., *Preservation of Electronic Formats: Electronic Formats for Preservation* (Fort Atkinson, Wis.: Highsmith, 1993), the proceedings of the 1992 WISPPR preservation conference (primarily concerned with medium preservation). Karen L. Hanus provides an extensive "Annotated Bibliography on Electronic Preservation" (p. 121-136).
21. National Digital Library Federation web pages describing the NDLF's initial steps and organization. (6/97)
22. National Preservation Office, National Library of Australia, Statement of Principles: Preservation of and Long-Term Access to Australian Digital Objects (last updated 17 Jan. 1997). A seemingly unexceptionable statement, except that it's wide-ranging coverage and national scope make it an important document; I don't think even the UK has gone this far in asserting the problem, the need and the means. (5/1/97)
23. Neavill, Gordon B., "Electronic Publishing, Libraries, and the Survival of Information," *Library Resources & Technical Services* 28:76-89 (Jan. 1984).
24. PADI (Preserving Access to Digital Information) is a full, up to date and very useful set of pages with resources, glossaries, information and links to bibliographies and other organizations, put up in Jan. 1997 by the Australian PADI Working Group (1/23/97).
25. Phillips, Margaret, Ensuring Long Term Access to Online Australian Publications: National Library of Australia Initiatives (January, 1997; paper at *Information Online & On Disc, Eighth Australasian Conference*, 21-23 January 1997, Sydney). A work group developed guidelines for developing a list of "Australian online publications which the Library would want to undertake to preserve.... we were not to consider how it could be done, only what ideally should be done." Discussions include policy and procedure development, capture strategies given the variety of publications, and access conditions once archived. References (links; some of which I've provided above).  
<URL:<http://www.nla.gov.au/nla/staffpaper/mphillips3.html>>

(4/22/97)

26. Preserving Digital Information: Final Report and Recommendations .  
Task Force on Digital Archiving, Commission on Preservation and Access and the Research Libraries Group. (5/20/96)

"At the end of 1994 the Commission on Preservation and Access (CPA) and RLG created a Task Force on Archiving of Digital Information charged with investigating and recommending means to ensure "continued access indefinitely into the future of records stored in digital electronic form." The 21-member task force, co-chaired with distinction by Donald Waters, Associate University Librarian, Yale University, and John Garrett, Chief Executive Officer of CyberVillages Corporation, recently completed their final report. RLG and CPA are making this widely available online and in print. Electronic versions are available from RLG's FTP server (ftp.rlg.org) and this Web site...."

"Copies of the printed, bound report are available for \$15.00 (prepayment required) from the Commission on Preservation and Access, 1400 16th Street, N.W., Suite 740, Washington, DC 20036-2217."

See also the companion commentary piece by Don Waters, below.

27. Van Bogart, John W. C., *Magnetic Tape Storage and Handling: A Guide for Libraries and Archives*(Washington, DC: Commission on Preservation and Access, June, 1995; available from the Commission at \$10: 1400 16th S. NW, Suite 740, Washington, DC 20036-2217), 34 pp. "This report was a joint effort of the Commission and the National Media Laboratory (NML), St. Paul, MN."
28. Waters, Donald J., The Implications of the Draft Report of the Task Force on Digital Archiving (that is, on Preserving Digital Information, above), given at the Fall, 1995 meeting of the Association of Research Libraries. Also reprinted in the *Newsletter* of the The Commission on Preservation and Access no. 85, January 1996, pp. [A-D] (2/24/96).

[....]

**APPENDIX II: PROJECT DIRECTOR VITA****PETER S. GRAHAM**

Associate University Librarian for Technical and Networked Information Services  
psgraham@rci.rutgers.edu; <URL:http://aultnis.rutgers.edu/pghome.html>.

**Education and Honors:**

M.L.S., Indiana University, 1981.  
D.Phil. student, Oxford University, England (Linacre College), 1973.  
B. Phil., Oxford University (Linacre College), 1972.  
M. A., Honors, Columbia University (New York), 1970.  
B. A., cum laude, Columbia University, 1969; Phi Beta Kappa.  
Univ. of Calif. (Berkeley), 1956-1958.

Winner, ALA / K.G. Saur Award for outstanding 1995 article published in *College & Research Libraries*, "Requirements for the Digital Research Library," 1996.

**Employment:**

Rutgers University Libraries: Associate University Librarian for Technical and Networked Information Services, 1987- .

Associate Vice President for Information Services, 1989-1992.

Selected University Activities at Rutgers:

1996 (July)-1997 (April), Acting Chair, Univ. Faculty Committee on Electronic Publication, Promotion and Tenure (member from 1995).  
1991, President's Committee on Computing and Information Planning.  
1991, President's Committee on Administrative Systems Efficiency.

Columbia University Libraries, New York City:

Assistant Director for Bibliographic Control, 1985-1987;  
Head, Book Acquisitions Department, 1981-1985;  
Acting Head, Serials Acquisitions Department, 1983.

Indiana University Libraries, Bloomington, Indiana:  
Systems Officer, 1979-1981.

Research Libraries Group, Branford, Connecticut:  
Senior Systems Analyst, 1975-1978.

**Governing Body Memberships:**

1995-1998 Council, Bibliographical Society of America; elected January, 1995.  
1993-2001 Council, American Library Association (ALA governing body; elected 1993; reelected 1997);  
1992- Member, Governing Board, Center for Electronic Texts in the Humanities (Princeton / Rutgers).

**Selected Activities:**

Follett Lecture Series, "Building the Digital Research Library: Preservation and Access at the Heart of Scholarship," Univ. of Leicester, England,

Mar. 19, 1997.

1997 Invited, funded participant, Fourth Dublin Core Metadata Workshop (OCLC / NCSA / NSF / Nat. Lib. of Australia), Canberra, Mar. 3-5.

Conference on Electronic Archiving and Preservation, keynote speaker (Univ. of Warwick, England: Nov. 27-28, 1995), British Library, Joint Information Systems Committee.

Cornell University / Research Libraries Group, Symposium on Digital Imaging Technology for Preservation, "Long-Term Intellectual Preservation" (Ithaca: March 17, 1994).

ALA/ALCTS Institute, The Electronic Library: Administrative Issues for Organization & Access, "Moving Beyond Traditional Electronic Services: Collection, Access and Text Management Issues" (Cincinnati: Oct. 17, 1993; repeated, San Antonio: Oct. 30, 1994).

Joint Harvard-MIT Workshop on Technological Strategies for the Protection of Intellectual Property in the Networked Multimedia Environment, "Intellectual Preservation and Electronic Intellectual Property," (Cambridge, Mass.: April 2-3, 1993).

**Selected Professional Associations and other activities:**

American Library Association, 1975- :

ALA Council (ALA governing body; elected 1993, 1997) 1993-2001;

Bibliographical Society of America (Council, governing body; elected 1995);

Founder and moderator, ExLibris, electronic bulletin board (listserv) for rare books and special collections community; 1650 subscribers, 1990-1995.

Member, Mid-Decade Planning Group, Research Libraries Group; developed and presented proposals to RLG Board, adopted Oct. 1993, 1992-1993.

Pierpont Morgan Library, New York: consultant on computerization; visits, analysis and report, 1987.

**Selected Publications:**

Review of E. B. Tribble, *Margins and Marginality* (Charlottesville: U. of Va., 1993), in *Renaissance Quarterly* (Winter, 49.4, 1996).

"Requirements for the Digital Research Library," *College & Research Libraries* 56 (July, 1995): 331-339.

*Intellectual Preservation: Electronic Preservation of the Third Kind* (Washington, DC: Commission on Preservation and Access, March, 1994), 8 pp.

"Quality in Cataloging: Making Distinctions," *Journal of Academic Librarianship* 16.4 (September, 1990), p. 213-218.

**APPENDIX III -- ACRONYMS:**

AAUP	Association of American University Presses
ARL	Association of Research Libraries
BOF	Birds of a feather (informal break-out groups)
CIMI	Consortium for the Computer Interchange of Museum Information
CLR	Council on Library Resources
CLIR	Council on Library and Information Resources, successor organization to the combined CLR/CPA from June, 1997
CNI	Coalition for Networked Information
CNRI	Coalition for National Research Initiatives
CPA	Commission on Preservation and Access
IETF	Internet Engineering Task Force
NINCH	National Initiative for a Networked Cultural Heritage
NSF	National Science Foundation
OCLC	Online Computer Library Center
PADI	Preserving Access to Digital Information [Australia]
RLG	Research Libraries Group
UKOLN	United Kingdom Office of Library and Information Networking