

"The Digital Library Federation - Goals and Priorities"

Remarks by Hans Rütimann, International Program Officer
Council on Library and Information Resources
Washington, D.C.

Background

The Council on Library and Information Resources (CLIR) is the outcome of the merger of the Council on Library Resources (CRL) and the Commission on Preservation and Access (CPA). CLIR's mission calls on the organization to identify the critical issues that affect the welfare and prospects of libraries and archives, to convene individuals and organizations in the best position to engage these issues, and to encourage institutions to work collaboratively to achieve and manage change. CLIR pursues its mission out of the conviction that information is a public good and of great social, intellectual, and cultural utility. It has set its sights on a few targeted programs: the Commission on Preservation and Access retains its identity as a program of CLIR, along with programs for Digital Libraries (the Digital Library Federation), the Economics of Information, and Leadership.

As a fundamental principle of all its programs, CLIR will encourage institutions to achieve and manage change through collaboration, and collaborative action is particularly important within the preservation community. The most striking evidence of this is the success of the ongoing effort to rescue through microfilming large portions of the deteriorating print-based collections in the U.S. and abroad. Since its inception, the Commission on Preservation and Access has worked to assure that knowledge produced by the scholarly communities of the world is saved and kept accessible and it will continue that role.

With advice from its standing committees and task forces, CLIR publishes materials that inform and instruct the preservation community, document the economic implications of establishing sound preservation environments for collections, frame the next set of issues to be considered within the changing definition of "preservation and access," and develop new strategies to sharpen the professional skills of individuals with preservation responsibilities. A few examples of recent publications indicate the range of our concerns:

- "Digitizing Historical Pictorial Collections for the Internet"
- "Preservation and Archives in Vietnam"
- "Digitization as a Method of Preservation?"
- "Mass Deacidification: An Update on Possibilities and Limitations"

The International Program

Because few of the critical issues of preservation and access today can be addressed without an international focus, CLIR maintains an International Program to help promote preservation awareness throughout the world. Through training seminars, workshops, translation projects, publications, and a policy of generous response to requests for counsel and advice from colleagues abroad, it continues to promote the long-term preservation and accessibility of information.

The International Program has focused its efforts to date on Eastern and Western Europe, the former Soviet Union, China, and Latin America. Thanks to support from The Andrew W. Mellon Foundation, the work in Latin America will be extended, and CLIR will initiate new activity in Southern Europe (for example, in Greece) and in South Africa. The Program also hopes to expand in Asia and is seeking funds to develop new projects there. Examples of activities that will be undertaken in the next several years include providing expertise for preservation-needs assessments in libraries and archives abroad; the development of cooperative filming and digitizing projects for specific collections; and designing strategies to increase the production and use of permanent paper.

Two specific examples:

The International Register of Microform Masters: An international register of microform masters is essential if

scholars and librarians are to know what has been filmed at locations throughout the world and to avoid duplication of effort. The Program encourages libraries and archives to contribute records to regional nodes for the collection, organization, and distribution of information about reformatted collections. It supports efforts to link these records to an emerging international register and to reach international consensus on the elements and the record structure for listings of digitized materials. A successful example is the European Register of Microform Masters (EROMM) where through ten partners in nine countries, some 40 European libraries have so far contributed more than 400,000 records of microfilmed items to the database. An exchange arrangement with the U.S. bibliographic network RLIN allowed the addition of another 1.9 million records to EROMM's database..

Another example is the recently concluded project "Translation and Dissemination of Preservation Knowledge in Brazil." Access to information often means translation of professional literature into other languages. An inter-institutional alliance of interested organizations in Brazil guided the project, which included the translation into Portuguese of 52 titles of preservation literature, from environmental control to digital conversion. The translations (plus videos) formed the basis for workshops throughout the country. In the process, the project's coordinators collected valuable information about the state of collections in more than 1,400 libraries and archives.

During a recent meeting (Aveiro, Portugal) with librarians and archivists from Lusophone countries, we offered to make all these materials available in Portugal, Portuguese-speaking countries in Africa, and Macau. Since access to professional literature on basic preservation concerns for print, sound, and images is a top priority in most countries, the offer was received with enthusiasm.

I mention support for traditional preservation efforts for print, images, and sound at some length since there is, particularly in developing countries, a danger of putting too much faith in digital solutions for the preservation of a country's heritage. Librarians and archivists in developing countries are often unaware that digital storage for long-term archiving of information requires careful planning, that many organizational, technical, and organizational issues have not yet been resolved, and that digitization is not a means for preservation unless a long-term plan assures the survival of digitally stored information. What is really disturbing is that in the general excitement about all things digital, many institutions have put on hold traditional and basic preservation activities.

Also, preservation concerns did not start with digital information. There is a long tradition of preserving cultural heritage. The recognition that much of the printed documentary record is in jeopardy because of the introduction of acidic paper around 1850 led to a widespread preservation movement, similar to the one we're beginning to witness for digital information. The preservation community, especially preservation managers and administrators, have much experience in selecting and preparing collections, and coordinating massive collaborative reformatting projects, mostly microfilming. Their experience is useful even though we're now dealing with an entirely new medium.

The Task Force on Archiving of Digital Information

An important contribution to the debate over long-term archiving of digital information came from the Task Force on Archiving of Digital Information, organized jointly by the Commission on Preservation and Access and the Research Libraries Group (RLG). The Task Force, in effect, was asked to report on ways that society should work with respect to the cultural record it is now creating in digital form -- not only through conversion from print to digital but, perhaps more important -- through a record that is "born digital." Consider that, by the year 2000, an estimated 75% of all U.S. Federal transactions will be handled electronically. When President Clinton leaves office, it is estimated that his administration will hand over eight million electronic files to the National Archives; and these administration files are but a minuscule portion of the electronic record generated by the corporate world. Most of this information is born digital and there is no printed record to fall back on.

The Task Force's 1996 report argues that "the problem of preserving digital information for the future is not only -- or even primarily -- a problem of fine tuning a narrow set of technical variables." Rather, "it is a problem of organizing ourselves over time and as a society to maneuver effectively in the digital landscape. It is a problem of building -- almost from scratch -- the various systematic supports, or deep infrastructure, that will enable us to tame our anxieties and move our cultural records naturally and confidently into the future" (Task

Force 1996:6)

The Task Force's report is available in full at the website of the Research Libraries Group (<http://www.rlg.org>). The Task Force was co-chaired by Donald J. Waters, now the Director of our Digital Library Federation.

Among the report's conclusions and recommendations are:

- The Task Force recognizes that most of the challenges associated with digital preservation are organizational, not technical.
- The first line of defense against loss of valuable digital information rests with the creators, providers, and owners of digital information.
- Certified digital archives must have the right and duty to exercise an aggressive rescue function as a fail-safe mechanism for preserving digital information that is in jeopardy of destruction, neglect, or abandonment by its current custodian.

The final report focuses on three essential questions: What does digital preservation entail? How do we organize ourselves to do it? What steps should we take to move forward? Time will not permit covering these questions in detail, but one major conclusion of the report cannot be emphasized enough: Our greatest challenges in the digital age are organizational rather than technical. Because we currently lack the infrastructure of practices, standards, and organizations needed to support preservation of digital information, the following elements of infrastructure must be considered:

- Legal bases for deposit and rescue. Nationally and internationally, legislation and agreements are needed to encourage legal deposit of electronic resources in archival repositories, to enable rescue of abandoned resources, and to facilitate access and use of archival files.
- Standards for description. Current library cataloging standards are not sufficient to describe access and contextual information about digital resources. Several efforts to address this issue are underway internationally. For example, existing registers of microform masters are examined for expansion to include digital items. In this context, see also the final report of the *RLG Working Group on Preservation Issues of Metadata* (<http://www.rlg.org/preserv/presmeta.html>).

There is more, and the need for sharing information about best practices across the wide spectrum of communities is overwhelming. Preservation of digital materials has emerged as a new, critically important field of interdisciplinary and international activity, and much work needs to be done.

[Introduction to the film "Into the Future"]

[Addressing our digital memory crisis begins with dialogue. CLIR prepared a discussion paper on digital preservation. Copies of the discussion paper are available, but first, if I may, I would like to show the documentary movie "Into the Future," on the preservation of knowledge in the electronic age. The movie was commissioned by the Commission on Preservation and Access and the American Council of Learned Societies. Funding was provided by the Alfred P. Sloan Foundation, the National Endowment for the Humanities, and the Xerox Corporation. It was shown on U.S. national television last January and created much discussion. The mainstream press picked up on the issue and several thoughtful articles appeared on the subject of digital preservation. The discussion paper raises the question "Why should we be concerned?" "Into the Future" provides several answers.]

Showing of the 1/2 hour version of "Into the Future."

After the showing: The movie sounds an alarm and raises questions; it does not provide answers. We hope that the answers will eventually be provided by groups such as represented at this conference, the Digital Library Federation (DLF) and its members, and other groups and individuals.

The Digital Library Federation (DLF)

(For this part, I'm relying extensively on information provided by Donald J. Waters, Director of the Digital Library Federation. For more information about the DLF, see CLIR's website (<http://www.clir.org/>).

The Council on Library and Information Resources is administrative home to the Digital Library Federation, which includes 19 members -- university research libraries, the Library of Congress, the National Archives, and the New York Public Library.

The primary mission of the DLF is to establish the necessary conditions for creating, maintaining, expanding, and preserving a distributed collection of digital materials accessible to scholars and a wider public. Participants in the Federation are committed to a shared investment in developing the infrastructure needed for libraries of digital works. The infrastructure is intended to enable digital libraries to bring together, or "federate," the works they manage for their readers.

The DLF has set the following program priorities:

- DLF will focus on libraries of materials born digital. It is critical that the library community moves from conversion -- except in specialized and well-justified cases -- to organization, access, and preservation of materials born digital. A high priority is developing the archival mechanisms that preserve the integrity and usability of these digital works over a long term. The Federation is increasingly turning the attention of libraries to the numerous -- and difficult -- issues associated with works born in, rather than converted to, digital form.
- DLF will help integrate digital materials into the fabric of academic life. A critical focus for such integration is to define the circumstances under which conversion to digital form is justified. Conversion projects that facilitate the extension of higher education and promise to improve the quality and lower the cost of research and education deserve special attention.
- DLF will help stimulate the development of a core digital library infrastructure. The highest priorities for attention at the present time are the network and systems requirements and means of authentication and authorization, the means of discovery and retrieval, and archiving.
- DLF will help define and develop the organizational support needed for effectively managing digital libraries. Organizational issues requiring early attention include identifying institutional values and strategies for managing intellectual property in digital form, and creating the conditions for the development of the professional skills needed for digital library management.

CLIR and DLF Plans for Digital Archiving

As one of its primary agenda items, CLIR and the DLF aim to ensure the persistence of digital information, both born digital or created by conversion. Much has been written about the need to document digital information in order for it to serve the role of "record" in an archival sense, as well as about the need to "migrate" digital information to new media in order to prevent loss from media decay and obsolescence. However, relatively little effort has gone toward answering the more fundamental question of how to ensure that digital documents will remain readable and understandable in the future.

At this time, one might usefully distinguish the following strategies:

- 1) Copying (no changes to the information)
- 2) Migration (adaptation to new hardware and software)
- 3) Emulation (new platforms mimicking previous platforms)
- 4) Archeology (doing nothing and trying to salvage neglected information when needed)

CLIR and the DLF commissioned Jeff Rothenberg, author of "Ensuring the Longevity of Digital Documents" (*Scientific American*, January 1995) to investigate approaches currently being considered to ensure the future accessibility and readability (i.e., longevity) of digital material. His preliminary findings challenge the concept of "migration." Since migration involves translation of each different type and format (text, images, sound,

video, animation, etc.), the process is highly labor-intensive. If a document is accepted for long-term storage based on an assumed migration strategy, Rothenberg argues, it is impossible to give even gross estimates of what will have to be done to it in the future, when this will have to be done, how much it will cost, or how long it will be before the document is corrupted by inappropriate translation.

In his further research, Rothenberg will concentrate on emulation as a solution to the problem of digital preservation, one that is predictable and testable and eliminates repeated translation of documents "which must inevitably lead to their corruption and loss." Others argue that "migration" is, at least for the shorter term, the only viable solution. Stay tuned.

Other activities of the DLF that relate to archiving are:

- Digital collections of licensed works. Libraries are pouring substantial resources into developing licenses for access to electronic journals, while still maintaining subscriptions to print copies of the same materials. Not all contracts provide for long-term access to the licensed materials, but terms in some licenses require the publisher to provide the library with a tape or CD-ROM containing a copy of the licensed work. The Copyright Division of the Library of Congress is developing the means to accept archival quality versions of copyrighted works deposited there, but project work is needed to ensure that the tapes and CD-ROMS deposited with libraries and archives as archival copies can and will serve the fail-safe purpose for which they are intended.

- Institutionalizing digital information. Much scholarly information in digital form is currently being managed as a cottage industry by individual faculty on behalf of a particular discipline. The directors of the DLF institutions are considering how DLF can provide leadership in bringing products emerging from this cottage industry into a more stable, institutional environment.

- Cornell has developed a proposal that DLF will support, to explore the requirements and means for preserving materials converted to digital form in a subset of genres.

- Extension of reach: *Making of America* Phase III. Conversion, as the Library of Congress has demonstrated, is a significant means of extending the reach of library collections in the service of general education. One of the founding goals of the DLF is to find ways to aggregate the existing digitized collections. This effort will focus primarily at the level of descriptive metadata as a means of integration and provide a testbed for exploring intersystem searching methods. There is another important dimension to the project: In several institutions (e.g., Yale and Cornell), large quantities of digitized Americana are inaccessible because the digital platforms on which the collections were built have become obsolete. *Making of America* Phase III thus affords the opportunity to develop and demonstrate migration techniques as a means of preservation of digital information.

Conclusion

Perhaps we should add a fifth strategy to the four already mentioned -- prayers. James Gleick reported in *The New York Times* ("The Digital Attic: An Archive of Everything," 12 April 1998) that "the Daiho Temple of Rinzai Zen Buddhism held a 'memorial service for lost information' in Kyoto and online." Gleick adds that "of course, the details are lovingly preserved, in English and Japanese, at its web site. A look at the web site reveals that there is much common ground between this conference and the Rinzai Zen Buddhists (<http://www.thezen.or.jp/jomoh/kuyo.html>):

"After the effort of transforming all this knowledge into electronic information has been completed, is it enough then to say that we are finished? And from there, can we truly make effective use of that which we have created? Sometimes, the answer is 'no.' To provide an example, there are many 'living' documents and softwares that are thoughtlessly discarded or erased without even a second thought. It is this thoughtlessness that has drawn the concern and attention of Head Priest Shokyu Ishiko. Head Priest Ishiko hopes that through holding an 'Information Service' and by teaching the words of Buddha, that this 'information void' will cease to exist."

We will all have to help him.