

Landscape of Digital Archiving*

Deanna B. Marcum**

Ms. Marcum describes some of the important projects currently underway in the area of digital archiving, covering federal programs and initiatives of non-profit organizations, universities, and international groups. She stresses that collaboration between librarians and publishers is needed to solve the preservation problem.

¶1 Much of the discussion about preservation today is focused on digital information. We take on this important problem, even though traditional preservation problems are many and are still very much with us. In some ways I think they are more critical, because they seem to have been put aside while we search for the funds to solve the looming question of preserving digital information. Many of the grant-funded projects for preservation microfilming have been reduced or eliminated, and much of the grant money that is available from private foundations and government seems to be directed toward digital archiving.

¶2 In this article I will briefly describe the kinds of programs that have been developed to address digital archiving. These solutions include federal programs and initiatives of nonprofit organizations, universities, and international groups. I will also address collaboration between publishers and librarians because, in the end, we are not going to solve the preservation problem unless we work with publishers.

Efforts to Develop Digital Archiving

Federal Programs

¶3 The federal initiatives in the area of digital archiving are focused primarily on infrastructure. Recently, the Library of Congress' (LC) National Digital Information Infrastructure and Preservation Program has received a lot of publicity because of the \$100 million that Congress appropriated to address some of the problems associated with digital preservation. The planning process has been completed, Congress has approved the plan that LC submitted,¹ and LC is now in the process of identifying collaborators—other institutions that will work with it to test certain principles

* © Deanna B. Marcum, 2004.

** Associate Librarian for Library Services, Library of Congress, Washington, D.C. This article is based on a paper presented at the conference "Preserving Legal Information for the 21st Century: Toward a National Agenda," Georgetown University Law Center, March 6–8, 2003.

1. LIBRARY OF CONGRESS, PRESERVING OUR DIGITAL HERITAGE: PLAN FOR THE NATIONAL DIGITAL INFORMATION INFRASTRUCTURE AND PRESERVATION PROGRAM (2002) [hereinafter NDIIPP PLAN], available at http://www.digitalpreservation.gov/repor/ndiipp_plan.pdf.

that have been identified in the plan. My organization,² the Council on Library and Information Resources (CLIR), worked with LC in doing several environmental scans to look at what is involved in preserving different types of content: electronic journals, e-books, digital sound recordings, digital video, digital television, and Web archiving.³

¶4 The National Archives and Records Administration is also working on an infrastructure project. In 2003, it had an extra \$22 million in the appropriations to begin addressing what is needed to build an electronic records archives. This is a very important project, given the need to keep government accountable through its records. In another federal program, the National Agricultural Library, working with land grant institutions and the Department of Agriculture, is taking a topical approach to identifying the important agricultural literature that needs to be preserved. Finally, there is the data storage research project that Victor McCrary and his researchers at the National Institute of Standards and Technology undertook with funding from CLIR. It resulted in a guide to the care and handling of optical media that CLIR has published.⁴

Initiatives of Nonprofit Organizations

¶5 With respect to preservation, we cannot begin to discuss the work of nonprofit organizations without mentioning the Andrew W. Mellon Foundation. It has done more than any other organization I know of to help us think through questions of both traditional preservation and digital archiving. Mellon funded seven institutions to do experiments in archiving electronic journals. Of these grants, two have been funded at the implementation level: the JSTOR initiative⁵ and Stanford's LOCKSS (Lots of Copies Keep Stuff Safe) project.⁶ The Mellon Foundation has made many institutional grants for digital archiving, often involving the creation of scholarly materials by the scholars themselves and the subsequent archiving of that work. Mellon continues to fund much of the research that is done in this area.

¶6 CLIR has aimed to educate the broader community about preservation issues through its publications and convening sessions. In collaboration with the Association of Research Libraries and three other organizations, CLIR recently

-
2. *Editor's Note*: The author served as president of the Council on Library and Information Resources, in Washington, D.C., from 1997 to 2003, when she was appointed associate librarian for library services at the Library of Congress.
 3. See *Building a National Strategy for Digital Preservation: Issues in Digital Media Archiving*, in 2 NDIIPP PLAN, *supra* note 1, at 13 app. 2, available at http://www.digitalpreservation.gov/report/ndiipp_appendix.pdf.
 4. FRED R. BYERS, CARE AND HANDLING OF CDs AND DVDs: A GUIDE FOR LIBRARIANS AND ARCHIVISTS (2003), available at <http://www.clir.org/pubs/reports/pub121/pub121.pdf>.
 5. See JSTOR—The Scholarly Journal Archive, at <http://www.jstor.org> (last visited June 28, 2004).
 6. See LOCKSS—Lots of Copies Keep Stuff Safe, at <http://lockss.stanford.edu> (last visited June 28, 2004).

issued a survey of the preservation landscape.⁷ It is a rather bleak assessment of how librarians think they are going to be handling the preservation work in the twenty-first century. I think they feel overwhelmed.

University Initiatives

¶7 Among the numerous university projects underway, I want to again mention LOCKSS. It is unlike anything else that is being done; it enables institutions to preserve material they do not own, but instead have licensed access. Indiana University and Emory University are testing this software, and it will be interesting to see how these two beta test sites perform.

International Initiatives

¶8 I want to mention a few relevant international initiatives. The UK-based Digital Preservation Coalition is a large group of stakeholders—libraries, archives, publishers, and creators—working together on the problem of digital preservation.⁸ They are trying to make the case to others that this is an important issue. They also are trying to coordinate their strategies so that what each of the stakeholders does will add up to something bigger than the sum of their individual efforts. They have been particularly successful in working in the areas of standards development and new services.

¶9 PADI (Preserving Access to Digital Information), an initiative of the National Library of Australia, is a very effective education forum that has managed to involve people from around the world both in experimental projects and in identifying literature that must be preserved, and then convincing people to take responsibility for it.⁹ Another international organization that doesn't have a lot of visibility because it is terribly underfunded is the European Commission on Preservation and Access.¹⁰ It brings together the national librarians of all of the European Union countries to discuss their collective digital preservation strategies. Its only practical experiment of which I am aware is a project to preserve images and make them available widely.

Collaboration between Publishers and Librarians

¶10 I would like to conclude with a few words about collaboration between publishers and librarians. In February 2002, CLIR and the Professional and Scholarly

-
7. ANNE R. KENNEY & DEIRDRE C. STAM, *THE STATE OF PRESERVATION PROGRAMS IN AMERICAN COLLEGE AND RESEARCH LIBRARIES: BUILDING A COMMON UNDERSTANDING AND ACTION AGENDA* (2002), available at <http://www.clir.org/pubs/reports/pub111/pub111.pdf>.
 8. See Digital Preservation Coalition, at <http://www.dpconline.org/graphics/> (last visited June 29, 2004).
 9. See PADI, Preserving Access to Digital Information, at <http://www.nla.gov.au/padi/> (last visited June 29, 2004).
 10. See European Commission on Preservation and Access, at <http://www.knaw.nl/ecpa/> (last visited June 29, 2004).

Publishing Division of the Association of American Publishers formed a joint working group of publishers and librarians. We decided it was so urgent to talk to each other that we agreed to put aside questions of copyright and journal pricing—issues that divide us. We focused instead on how digital technology is changing the very nature of our two enterprises and the implications of those changes. We identified nine things, excluding copyright and pricing, that we needed to talk about, and digital archiving was at the top of the list. We have taken as our challenge to answer the question, for publishers and librarians alike, “What constitutes a preservable digital document?” It is amazing how hard that is. Publishers keep saying they will be happy to produce a preservable digital document if we can tell them what it is, and we librarians have a very difficult time saying what it will look like. But we continue to make progress. The best news of all is that we have had enough success talking about this and the eight other issues identified in our last meeting that a publisher from one of the very big publishing houses said he thought we had enough experience and trust now that we can bring copyright back to the table. We take that as great progress; if we don’t find a way to work with publishers, we’re not going to be able to preserve materials because we don’t own much these days.

¶11 Finally, in 2002, the International Federation of Library Associations and Institutions (IFLA) and the International Publishing Association issued an important joint statement on archiving and preservation of digital information.¹¹ They list seven principles in this document and one of the strongest statements, in my view, is that libraries should be the focus of digital archiving through arrangements with publishers; the responsibility should not be turned over to the publishers.

¶12 These are just a few of the important projects that are underway in the area of digital archiving.

11. Int’l Fed’n of Library Ass’ns & Insts. & Int’l Publ’g Ass’n, *Preserving the Memory of the World in Perpetuity: A Joint Statement on the Archiving and Preserving of Digital Information* (adopted June 27, 2002), available at <http://www.ifla.org/V/press/iflaipa02.htm>.