

**“The DODL, the NDIIPP, and the Copyright Conundrum”
Friday Keynote Address to Fiesole 2004
The Fiesole Collection Development Retreat Series
The European University Institute (EUI)
Villa La Fonte, Florence, Italy
19 March 2004**

**Deanna B. Marcum
Associate Librarian for Library Services
Library of Congress
Washington, D.C., U.S.A.**

I feel deeply privileged to be here for three reasons.

One: it is always a privilege to be in Italy.

Second: it is a privilege to be on the same program with so many distinguished speakers. They include my friend Michael Keller, who spoke so impressively last night.

And third: it is a privilege to speak at a meeting attended by both publishers and librarians. It gives me an opportunity to say some things that I want urgently to address to both. In what follows, I hope to explain not only what those things are but also why they are urgent.

1

Currently, in the United States, I am involved in two programs that are of particular importance to the Library of Congress and potentially of importance to all of us who care about maximizing the benefits of digital technology for information service. Both programs have acronymic names that are, at best, mysterious, and at worst, weird.

One is called *the DODL* and the other is *the NDIIPP*. DODL stands for Distributed Open Digital Library and NDIIPP stands for National Digital Information Infrastructure and Preservation Program.

Some of you have acquaintance with the DODL because your institutions are members of the Digital Library Federation, which is creating it. Mike Keller certainly knows about it; as chair of the Steering Committee of the Digital Library Federation, he has been trying to get the federation to provide the DODL with some more appealing name. It pleases me to be able to say here, by the way, that the DODL's development has become international because the Digital Library Federation has just been joined by the British Library, the federation's first strategic partner outside of the United States, and others may follow.

Within the United States, the Digital Library Federation has thirty-two member institutions and four affiliates. Most of the institutions are the libraries of major research universities but the federation also includes the New York Public Library, the Council on Library and Information Resources, the U.S. National Archives, and the Library of Congress.

Last fall, the members of the federation committed themselves to creating jointly something first envisioned in 1995 at the time of the federation's founding—a collaborative digital library that will provide global electronic access to collections in multiple institutions. This Distributed Open Digital Library—the DODL—will provide users with one point of entry to multiple, digitized collections. Moreover, the federation hopes to provide services using this distributed library that may include a deep finding system and specialized portals to help scholars locate material for their individual fields

and specific studies. The DODL's collection will begin with materials in the humanities and social sciences, and will transcend regions to concentrate on topics, themes, genres, and formats, including material in special collections.

The Digital Library Federation, which is supported primarily by its members, including the Library of Congress, is in the process of generating financing for the DODL, of appointing a coordinator for it, of forming a collections-development working group to plan development of the DODL's content, and of forming a technical working group that will develop an enabling infrastructure for the collaborative library. As Mike Keller explained when he announced these things on behalf of the federation last fall, the DODL will build on a range of achievements in digital library development by federation members and others internationally over the past several years. The DODL's time has come.

At the same time, the National Digital Information Infrastructure and Preservation Program—the NDIIPP—is under development at the Library of Congress. This program is focused on *preserving* the kinds of materials that digital libraries such as the DODL are creating.

Many of you, like me, are concerned about how to preserve, long-term, all the digital materials that publishers, libraries, scholars, and many others are creating. Digital media are much less stable than paper. Obsolescence in rapidly changing computer hardware and software can render digital materials unreadable. And new digital formats are being created more rapidly than we are creating capabilities for preserving them.

This concern has reached the Congress of the United States, which has appropriated one hundred million dollars for the NDIIPP's development. Our Office of

Strategic Initiatives at the Library of Congress has developed a plan for the NDIIPP that drew widely on expertise from many kinds of organizations outside the library, and the plan calls for the NDIIPP to be developed as a partnership with other institutions. This winter the library has been collaborating with the National Science Foundation in the U.S. to fund research for the NDIIPP through a call for proposals that will result in cooperative agreements with partnering institutions.

You will notice that I have stressed the collaborative nature of both the DODL and the NDIIPP. That is because neither the Library of Congress nor any other one institution, in the United States or elsewhere, can, by itself, hope to create an electronic gateway to the digital resources of libraries around the world and to overcome all the challenges to preserving these resources. Like it or not, the future development of library service depends on collaboration.

However, there is something more towards which my discussion of the DODL and the NDIIPP is headed, even though I feel slightly silly every time I use these terms. There is nothing silly, however, about the visions that these programs have of where they are headed. The following is a quote from the NDIIPP Plan:

The vision of the National Digital Information Infrastructure and Preservation Program is to ensure access over time to a rich body of digital content through the establishment of a national network of committed partners, collaborating in a digital preservation architecture with defined roles and responsibilities.¹

Please keep this in mind while I now add the following vision statement for the DODL. I owe this statement to Mike Keller, who, as chairman of the Steering Committee

of the Digital Library Federation, has described what the DODL is envisioned to make possible. Here is what he sees by the year 2013:

Readers and researchers, students and teachers around the world search for words and phrases using Internet engines that perform federated and broadcast searches on multiple virtual memories. More than fifteen million books have been digitized and conversion completed of the backset of virtually all scholarly journals. Government documents going back to the eighteenth century have been converted to digital form as well. All of these texts are available as page images and, for most of them, there are versions coded using XML available for searching. All books, journals, and other documents published since the turn of the twenty-first century are available in digital form as well. Words, phrases, and even concepts can be searched in the virtual global collection as a result.²

What a wonderful prospect these two vision statements hold out. Vast quantities of recorded information and knowledge will be easy to access by scholars, students, and other researchers around the world, and will be safely preserved for access by generations of researchers yet to come! Think how empowering this is going to be for all of us! And it is all now becoming technologically possible.

Alas, it is not going to happen. That is, it is not going to happen anything like as fully as it could. Not without another kind of collaboration—a collaboration that finds ways to reconcile global library access with individual intellectual property rights.

The Library of Congress contains the Copyright Office of the United States, a division separate from the one that I head. Nothing that I am about to say should be taken to represent the Copyright Office. It has enough challenges without having to take responsibility for opinions that are strictly mine.

By statute, the Copyright Office administers United States copyright laws. Its duties include giving advice about copyright law to Congressional legislators and also to U.S. government officials preparing cases before courts. And as some of you know, officials of the Copyright Office take part in U.S. delegations to international bodies. These include the World Intellectual Property Organization and the World Trade Organization's Council on Trade-Related Aspects of Intellectual Property Rights, well known by the acronyms TRIPS and WIPO. (I don't know why everything I am talking about today is so bizarrely named!)

Our Copyright Office in the U.S. can make recommendations concerning the interpretation of copyright law, and it has been called upon to provide mediation services. For example, it moderated negotiations between intellectual property owners and public universities to determine the liability of states of the United States for violations of federal intellectual property laws. And it helped bring about a compromise among affected parties concerning copyright considerations in digital distance education. It has done all these things, in its own words, "to ensure that the nation maintained a strong and effective copyright system—one that served both owners and users of copyrighted works."³ That recognition of and respect for multiple and sometimes competing interests is important.

Also of great importance is the responsibility of the Copyright Office for registering copyright claims. Here is where the magnitude of libraries' copyright concerns becomes dramatically visible. In the 2003 fiscal year for the Library of Congress, the Copyright Office received 526,138 claims to copyright covering more than 800,000 works. Eight hundred thousand in one year alone!

Under mandatory deposit provisions of U.S. copyright law, copyright owners must deposit two copies of any work published in the U.S. with the Copyright Office. These deposit copies go to the Library of Congress, which may add them to its collection or give them to (or exchange them with) any other library. But this deposit requirement pertains to physical copies, not to underlying rights. Concerning digital copies, the Library of Congress makes some deposited works available for use electronically on a local area network by agreement with copyright owners. But beyond that, an expert commissioned to assess copyright laws for the NDIIPP has concluded, "Nothing in the current law would permit LC to make deposit copies generally available in digital form on a publicly accessible network."⁴

So much for digitizing and providing worldwide Internet access, now and in perpetuity, through the DODL and the NDIIPP to the 800,000 works a year for which U.S. copyrights are sought, or whatever portion the Copyright Office registers and the Library of Congress takes in.

Now before anybody here starts throwing pizza at me for seemingly denigrating copyright, please wait to see whether I actually do so. And please remember the concern of the Copyright Office to serve "both owners and users of copyrighted works."

In actuality, libraries already are providing Internet access to large quantities of digitized and digitally created material. More than 7.8 million items are currently available electronically from our “American Memory” collection and others on the Library of Congress Web site. Free access is available to the texts of some ten thousand books available online from the Electronic Text Center of the University of Virginia. The University of Michigan’s digital library offers access to fifty-seven different electronic collections containing some 5.8 million pages of text. The New York Public Library and the libraries at Cornell, Harvard, and many other universities also have put thousands of texts, images, and other resource materials online. But except in some instances governed by special considerations or agreements, all of this material is in the public domain. That means that most of it predates the twentieth century, because much of what has been published in the past century and this one remains under copyright.

In the print world, libraries simply bought books, journals, and other materials, which the “first sale” provisions of U.S. copyright law enabled them to provide in (I quote) a “public distribution” to whomever, however they wished. But limits were built in. That is, if you couldn’t come physically to a library, you couldn’t get the library book you wanted, except possibly through the relatively laborious process known as interlibrary loan. So each library bought a lot of the same books and journals that other libraries did, which made libraries a good market for publishers.

In the electronic world, however, anybody with a computer properly programmed and Internet-connected can get a digital book from anywhere without leaving an office or home. So publishers generally lease, not sell, current journals and other research materials in digital formats to libraries, under agreements that restrict access to individual

libraries' patrons. Again in the words of the copyright law expert for the NDIIPP, "Making copies of a work available for public downloading over an electronic network qualifies as a public distribution. However, neither the courts nor the Copyright Office has yet endorsed a 'digital first sale doctrine' to allow users to retransmit digital copies over the Internet."⁵

This means that libraries cannot digitize and provide unrestricted Internet access even to works that they already own if such works remain under copyright. And on top of that restriction on digital access, the length of time to which copyright can be extended in the United States keeps going up and up.

Copyright is enshrined in the U.S. Constitution. Article 1, Section 8, provides (and I quote): "The Congress shall have power to promote the progress of science and useful arts by securing for limited times to authors and inventors the exclusive right to their respective writings and discoveries." In 1790, in the nation's first copyright act, the Congress decided that the appropriate limit for copyright holding should be fourteen years. But just in the past forty years, Congress has extended the term for copyright eleven times. The most recent was in October of 1998, when the Congress determined that the limit should be the length of an author's lifetime plus *seventy* years. And in 2003, the Supreme Court of the United States upheld that legislation by ruling the extension allowable under the Constitutionally granted Congressional power.

In short, if someone now publishes something at the age of forty and lives to be seventy, that author and the author's heirs can keep it under copyright, unavailable or under restriction for Internet access, for a full century. If the author publishes earlier and dies later, well, rightly or wrongly, the open contents of the Distributed Open Digital

Library will be diminished accordingly. And so will the open content of all other digital libraries.

3

“Wrongly,” many librarians and their professional organizations have complained. The Copyright Extension Act of 1998 is one of several developments by which they have felt increasingly restricted. The Digital Millennium Copyright Act, passed about the same time, and the Technology, Education, and Copyright Harmonization Act, passed in 2002, also seem to librarians to confuse, circumscribe, or both the rules for “fair use,” interlibrary loan, instructional resource use, patron privacy, and library preservation. Concerns about all this have been formally expressed to the U.S. Copyright Office and elsewhere by the American Library Association, the American Association of Law Libraries, the Association of Research Libraries, the Medical Library Association, and the Special Libraries Association.⁶ Moreover, in the words of the executive director of the Association of Research Libraries, as digital resources increase, “contractual licenses are supplanting copyright laws, with content owners mandating more restrictions on who uses resources and how these resources may be used.”⁷

In other words, our society may be responding to technology’s growing ability to expand access by increasing legal restrictions on it.

While the DODL is dealing with that challenge, the NDIIPP gets to deal with related concerns about digital preservation. Libraries fear that their preservation function is becoming increasingly difficult because they can only lease, not buy, most digital publications. Licenses for such publications often prohibit copying for archival purposes.

Many works, therefore, exist only on the servers of content providers, where they may be subject to corruption of various kinds and preserved only selectively. Publishers have little incentive to invest in preserving works with diminishing commercial value. Also, if publishers go out of business, the content they have controlled could be abandoned. How many digital creations of scholars and other writers are going to be available to students and other researchers in the future? The NDIIPP could be stymied by legal restrictions no matter how brilliantly its partners may meet preservation's technological challenges.

Intellectual property restrictions have become of such concern that some copyright scholars have called them "the second enclosure movement."⁸ Among others, James Boyle of the Duke University Law School compares today's restricting of the intellectual public domain to the famous movement in English economic history by which large landowners began fencing lands called "commons," which once had been open to all. Economic historians still debate whether the agricultural efficiencies of that enclosure sufficiently offset the social disruptions it produced, but Boyle heartily protests what he sees now as "the enclosure of the intangible commons of the mind"⁹ by the extension of copyright and patent laws. To illustrate the extent to which libraries are intimidated, Boyle likes to tell about a search he once made concerning a poem on a library Web site. When he inquired about the poem's source, the library quickly responded with a promise to remove the poem from distribution. The library assumed he was planning to sue for violation of copyright.

As some members of this audience already are probably feeling, it is unfair of me to continue as if publishers are the bad guys in all this, the greedy graspers of exclusionary privilege. But to paraphrase a famous observation, I have looked at the copyright holders, and they are us. Who among us has not published a copyrighted article or book? The ones I have published are far from making me rich. But among other things, copyright encourages acknowledgement of authorship, discourages plagiarism, and prohibits the theft of saleable property. Nobody, including Professor Boyle, disagrees about that.

Moreover, the *extension* of copyright protection to which I referred was chiefly supported by commercial interests other than publishers. The Association of American Publishers, representing a wide range and large number of publishing houses, “neither supported nor opposed” the legislation because of “member disagreements,” in the association’s words, “over the desirability of term extension as a matter of public policy.”¹⁰

Like the rest of us, of course, publishers do act in opposition to threats to their livelihood. International piracy is among such threats. The Association of American Publishers has supported legislation designed, in the words of Patricia Schroeder, president and chief executive officer of the association, to “ensure effective copyright protection and enforcement in the digital environment and global marketplace”¹¹ The publishers’ association is conducting a campaign to curb copyright infringement internationally by promoting copyright recognition and taking legal action against violators. Last year, in fact, the Association of American Publishers cooperated with local authorities to raid photocopy shops in Malaysia and seize what the association charged

were infringing copies of textbooks originally published by five American publishing houses.

Publishers also are represented in the International Intellectual Property Alliance, along with many other kinds of U.S. companies in what is called the “copyright community.” The alliance estimates that losses from piracy of U.S. copyrighted materials by people using inexpensive reproduction technologies exceed twenty billion dollars annually, not including losses through Internet piracy.

Large publishers in particular are additionally worried about Internet distribution of their publications within the U.S. For example, some small publishers have allowed libraries to lend their books electronically, but as Laurence Kirshbaum, chairman of the books division of AOL Time Warner, has been quoted as saying, “There is an inherent danger that would worry me—you are opening yourself up to being copied wildly without control.”¹² The publishing association’s president, Pat Schroeder, has similarly cautioned:

*We want free expression and no censorship but this is very different from saying that everything on the Net should be free I know that this is a very appealing mantra. The only problem is that after you give away all the content, it’s very doubtful that new, high quality content will appear to take its place because the people producing it have mortgages. If they aren’t going to be paid, they’ll look for work elsewhere.*¹³

At the same time, however, President Schroeder has staunchly supported librarianship in the digital information era. “Those who declared librarians obsolete when the Internet rage first appeared are now red-faced,” she has said.

“The Internet is full of ‘stuff’ but its value and readability is often questionable People found out that they needed librarians more than ever to navigate the sea of material out there.”¹⁴

The fact is that publishers and librarians, rather than being enemies, are close collaborators. As President Schroeder has also observed, they have worked together for years to promote books, reading, and education, and to protect free expression. They continue to collaborate with frequency, particularly in defense of free speech and other Constitutional rights that both uphold. In this very month, the Association of American Publishers applauded creation of a campaign to reduce the intrusion of anti-terrorism legislation into the privacy of library circulation and bookstore purchase records. In February a year ago, the Association of American Publishers and other organizations representing publishers, bookseller, journalists, and authors supported a court action that the American Library Association brought to protect Constitutional rights to information access for library patrons. These are but two of many examples of unity between publishers and librarians in promoting and defending rights to expression and access.

Now, urgently, we need, in another way, to extend that collaboration.

5

We need a collaboration of libraries and publishers, along with scholars and other authors and creators, to ensure that the DODL can create for the world, and also that the NDIIPP can preserve for posterity, an accessible assemblage of digital resources—an assemblage that isn’t always a century or more behind the growth of the world’s

knowledge, or so restricted that only a privileged class of users can gain access to it. Yet at the same time, we must retain the right for authors and inventors to benefit from their work that the U.S. Constitution says is needed “to promote the progress of science and useful arts,” and that Pat Schroeder says is needed to enable creators to support themselves.

Please observe that this is not a simple matter of librarians versus publishers, or of knowledge users versus knowledge creators. Authors who need copyright to protect intellectual property rights in works of scholarship will be able to produce new works only if not prevented by intellectual property restrictions from access to previous works of scholarship. Publishers who see long-term commercial value in their digital products may preserve them in libraries of their own or of others, just as libraries that help scholars and teachers create and disseminate digital materials may become those materials’ electronic publishers. These are multiple and overlapping needs to sort out and satisfy. Can we do it together?

In fact, can we build on collaborations already occurring?

Can we build on the search for mutual understanding that we began in December of 2001 when the Council on Library and Information Resources and the Scholarly Publishing Division of the Association of American Publishers joined in support of a Working Group of publishers and librarians to address issues of common concern? The working group’s mission statement acknowledges the profound effect of digital technology on both groups and recognizes that librarians and publishers share many goals and face similar problems.

Can we build on the creation of JSTOR, which stands for journal storage (at last, an acronym that makes sense!) JSTOR preserves and provides online access to digital copies of some 390 scholarly journals. It does this by licenses with publishers to maintain a “moving wall” of three to five years, in most cases, which is the time between the most recent issue published by a journal and the most recent issue available from JSTOR. Through this arrangement, as JSTOR puts it, JSTOR can “avoid jeopardizing publishers’ subscriptions and revenue opportunities from current and recent material while also enabling libraries and researchers to rely on JSTOR as a trusted archive, providing both preservation and access This balancing of interests among publishers, libraries, and scholars is at the center” of JSTOR’s approach.¹⁵

Can we build on other efforts, funded like JSTOR by the Mellon Foundation, in which publishers and the libraries of several major, individual universities have been exploring archiving possibilities for digital materials? For example, can we build on the program led by Stanford University called LOCKSS? There’s another sensible acronym: it stands for Lots Of Copies Keep Stuff Safe. Through agreements with publishers, multiple universities collaborating in LOCKSS keep digital materials preserved for posterity through a computerized system of mutually checking and replenishing repositories.

Can we build on electronic services created by publishers with library partners to improve access to biomedical journals? Some services in which publishers have invested heavily to deliver information directly to the desktops of healthcare professionals also provide such information through libraries, under license or via free access. For example, publishers supply biomedical article data, with links to full texts, to a service of the U.S.

National Library of Medicine called PubMed, which also makes approximately 100 journals available without charge within one year of publication.

Can we build on ideas for new kinds of arrangements, whereby, for example, scholars give their universities nonexclusive rights to make their output electronically accessible in “pre-publication” forms while remaining free to publish formally and commercially the portion of their research for which there is a market—an idea suggested by Malcolm Litchfield of the Ohio State University Press?¹⁶ Or is there some way that libraries can trade marketing opportunities to publishers in exchange for greater latitude in providing access to works under copyright?

Can we build on findings in the testimony, giving during strategic planning for the NDIIPP, that if income streams could be protected, many content firms would welcome digital preservation of their products in trustworthy libraries?

I don't have specific answers to all these questions. But I see lots of arenas in which people of good will—publishers, librarians, content creators, and information consumers—are trying to accommodate each other's needs. And on these efforts, I want to encourage all of us to build.

It is urgent that we do so because already we are developing the DODL and the NDIIPP and many other wonderful and ingenious means, internationally, of taking full advantage of the digital era's new information technologies. We can not let them bog down in legal disputes and operational confusions about intellectual property rights. We must not.

In closing, I must admit that my own institution, the Library of Congress, has not always been the most eager collaborator in the world. The vastness of our collections, and

our standing as the closest thing that the United States has to a national library, may sometimes have made us feel impervious to concerns within the rest of the information community. But if that notion ever made sense, which I doubt, the digital era makes it wholly untenable. In my role as head of Library Services at the Library of Congress, I am eager to work with others at home and around the world on copyright among many other issues. Thank you for this opportunity to say so.

#

¹ National Digital Information Infrastructure and Preservation Program, *Preserving Our Digital Heritage: Plan for the National Digital Information Infrastructure and Preservation Program, A Collaborative Initiative of the Library of Congress* (Washington, D.C.: Library of Congress, 2002), pg. 65; available in print and at www.digitalpreservation.gov.

² Michael Keller, "A Rose Is a Rose by Any Other Name; What's a DODL?" Address to the Digital Library Federation Fall Forum, Albuquerque, New Mexico, Nov. 18, 2003; available at http://www.diglib.org/forums/fall2003/keller/DLF-keller-031118_files/slide0019.htm.

³ *Annual Report of the Librarian of Congress for the Fiscal Year Ending September 30, 2002* (Washington, DC: Library of Congress, 2003), pg. 31.

⁴ June M. Besek, *Copyright Issues Relevant to the Creation of a Digital Archives: A Preliminary Assessment* (Washington, DC: Council on Library and Information Resources and Library of Congress, 2003), pp. 7-8.

⁵ *Ibid.*, pg. 3.

⁶ “Comments of the Library Associations Before the Library of Congress, the United States Copyright Office and the Department of Commerce, National Telecommunications and Information Administration,” Washington, D.C., Aug. 4, 2000; available at <http://www.arl.org/info/letters/dmca%5F80400.html>.

⁷ Duane Webster, “The Practical Realities of the New Copyright Laws: A Librarian’s Perspective,” address to the Modern Language Association Conference, New York City, Dec. 28, 2002, pg. 10.

⁸ James Boyle, “The Second Enclosure Movement and the Construction of the Public Domain,” 2003; available at <http://www.creativecommons.org/licenses/by-sa/1.0> and <http://www.law.duke.edu/journals/66LCPBoyle>.

⁹ *Ibid.*

¹⁰ “Association of American Publishers Releases Statement on Eldred Decision,” press release by AAP, Washington, D.C., Jan. 15, 2003; available at <http://www.publishers.org/press/releases.cfm?PressReleaseArticleID=119>.

¹¹ “Publishers Applaud Bipartisan Commitment to Effective Copyright Protections,” press release by the Association of American Publishers, Washington, D.C., July 17, 2003; available at <http://www.publishers.org/press/releases.cfm?PressReleaseArticleID=163>.

¹² David D. Kirkpatrick, “Battle Over Access to Online Books,” *New York Times*, June 17, 2002; available at <http://www.nytimes.com/2002/06/17/technology/17LIBR.html>.

¹³ “Publishers, Librarians Share Vital Role in Digital Information Age, Schroeder Tells Australian Librarians,” press release by the Association of American Publishers, May 22, 2002; available at <http://www.publishers.org/press/releases.cfm?PressReleaseArticleID=13>.

¹⁴ *Ibid.*

¹⁵ “JSTOR: The Moving Wall,” JSTOR Web site text: <http://www.jstor.org/about/movingwall.html>.

¹⁶ Malcolm Litchfield, “. . . But Presses Must Stress Ideas, Not Markets,” *The Chronicle of Higher Education*, June 28, 2002, pp. B 9-10.