Issues in the Preservation of Born-digital Scholarly Communications in Law*

Richard A. Danner**

Professor Danner provides a general introduction to some of the impacts of "born-digital" information on scholarly communications, then discusses the possible impacts on scholarly communications in law, focusing on questions involving the preservation of born-digital information.

New Models for Scholarly Communication

¶1 In the traditional print-dominated publishing environment, journals and monographs provided the primary means for dissemination of scholarly information while serving as artifacts that could be used by libraries or other institutions to preserve the knowledge created. With the emergence of an electronic publishing environment characterized by the Internet's powerful and immediate impacts on communication of research results, less attention was paid initially to preservation of scholarship published in electronic media than to the benefits those media provide for disseminating it.[1]

¶2 The processes of scholarly communication involve more than formal publication and library-centered research. Every field of scholarly activity employs both formal forms of publication and dissemination of research results (much of which is acquired by and preserved in libraries), and less formal forms of communication (letters, conversations at meetings, telephone conversations, e-mail, circulation of draft papers, etc.).[2] Disciplines differ in the extent to which their

---

** Rufty Research Professor of Law, and Senior Associate Dean for Information Services, Duke University School of Law, Durham, North Carolina. 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. It has also benefitted from comments of participants in workshops at the Duke Law School on December 13, 2001, at the University at Buffalo Law School on March 15, 2002, at the Yale Law Library on February 15, 2002, and at the University of Cape Town Chancellor Oppenheimer Library on September 26, 2003.

1. Although there is now much more discussion about preserving digital information, preservation of paper information sources remains a difficult problem for libraries. *See* Robert C. Berring, *Legal Information and the Search for Cognitive Authority*, 88 CAL. L. REV. 1673, 1684, n. 29 (2000) ("One of the sad failures of librarianship has been the inability to develop reasonably priced means of preserving books."); Kevin M. Guthrie, *Archiving in the Digital Age*, EDUCAUSE REV., Nov./Dec. 2001, at 57, 58 ("One does not have to spend much time in a large library to find paper volumes and documents that cannot be used for much longer.").

2. For a description of informal communication patterns in the sciences, see Rob Kling & Lisa Covi, *Electronic Journals and Legitimate Media in the Systems of Scholarly Communication*, 11 INFO. SOC'Y 261, 264 (1995) ("Scholars . . . often learn about new studies and results in their immediate areas well before they are published—through collegial conversation, conference presentations, attending invited seminars, acting as journal editors and reviewers, and receiving manuscript drafts or preprints from close colleagues. Active scholars are usually well positioned in these (primarily) verbal networks.").

members rely on particular forms of informal communications, but in all fields, there is much more to scholarly communication than the formal publications traditionally collected by libraries.

¶3 As scholars increasingly employ electronic media to create and to distribute both formal and informal forms of scholarly communications, the distinctions between formal and informal communications have become increasingly blurred. Draft versions of scholarly papers (original or revised after presentation and comment) are "published" and disseminated in the same digital media as final and definitive versions, and can remain accessible well after a "final" version is published. Comments on the papers, too, can persist on the Internet in e-mail messages or other forms of electronic communications long after they are first offered.[3] Many scholars have embraced the possibilities of this suddenly much richer information environment,[4] in which both traditional forms of published knowledge and less formal (and previously less accessible) forms of information can now be disseminated in common digital formats. For those concerned about organizing, preserving, and making accessible the scholarly literature, however, the convergence of informal and formal scholarly communications brings to the fore a number of issues (version control, persistent access, and retrievability) that were either of less concern or handled differently when final print publication was all that mattered for most purposes. In many ways, the increased accessibility and persistence of informal scholarly communications has created a new form of *grey literature*.[5]

¶4 This article provides a general introduction to some of the impacts of *born-digital*[6] information on scholarly communications, then discusses the possible

3. In some disciplines, journals have traditionally included letters or other forms of commentary on published research. These materials have persisted, even in print, but may not be known or accessible to readers of the original articles. *See* Stephen Adams, *Information Quality, Liability, and Corrections*, ONLINE MAG., Sept./Oct. 2003, at 16, 21.

4. *See, e.g.*, Joost Kircz, *New Practices for Electronic Publishing 1: Will the Scientific Paper Keep its Form?* 14 LEARNED PUBLISHING 265, 267 (2001) ("The fascinating issue is that in an electronic publishing environment we are indeed able to integrate these various formal and non-formal means of communication. . . .").

5. Grey literature is typically thought of in terms of "non-conventional, fugitive, and sometimes ephemeral publications" outside the control of commercial publishers. N.Y. ACAD. OF MED. LIBRARY, WHAT IS GREY LITERATURE, *at* http://www.nyam.org/library/greywhat.shtml (last visited July 5, 2004). A more appropriate definition for the current environment might be "Grey literature is really a type of informal communication, which on a scale of formality, fits in somewhere between conversation and normal publication." EUROPEAN ASS'N FOR GREY LITERATURE IN EUR., THE VALUE OF GREY LITERATURE, *at* http://www.kb.nl/infolev/eagle/what_is_gl.htm (last visited July 5, 2004). *See generally* Brian S. Matthews, *Grey Literature: Sources for Locating Unpublished Research*, 65 C. & RES. LIBR. NEWS 125 (2004).

6. The term *born-digital* will be used here simply to encompass material that was "created in digital form rather than converted from analog to digital." Amy Friedlander, *Summary of Findings, in* COUNCIL ON LIBRARY AND INFO. RES. & LIBRARY OF CONGRESS, BUILDING A NATIONAL STRATEGY FOR DIGITAL PRESERVATION: ISSUES IN DIGITAL MEDIA ARCHIVING 2 (2002), *available at* http://www.clir.org/pubs/reports/pub106/pub106.pdf. Born-digital is also sometimes used more restrictively to include only "materials that are not intended to have an analogue equivalent, either as the originating source or as a result of conversion to analogue form." NEIL BEAGRIE & MAGGIE JONES, PRESERVATION MANAGEMENT OF DIGITAL MATERIALS: A HANDBOOK §1.3 (2001), *available at* http://www.dpconline.org/text/intro/definitions.html.

impacts on scholarly communications in law, focusing on questions involving the preservation of born-digital information.

Electronic Archiving of Scholarly Communications

¶5 In some fields, scholars and researchers have responded to rising journal prices and other problems in their own scholarly communication systems by establishing electronic archives of articles or papers in *eprint* form,[7] frequently with the goal of reporting recent research prior to its eventual appearance in the journals of the discipline.[8] Although many of the papers posted directly by authors into Internet archives or repositories are later published formally in journals (sometimes after revisions prompted by readers' comments), other papers remain available solely in eprint archives. Proponents of electronic repositories aim to create more efficient scholarly communication systems within individual disciplines, as well as the means to disseminate scholarship to wider audiences. "Open archives" advocates, such as physicist Paul Ginsparg,[9] cognitive scientist Stevan Harnad,[10] and philosophy professor Peter Suber,[11] are devoted to using these technologies to change the structures of scholarly communication. "The fundamental idea [of the open archives movement] is that authors would deposit preprints and/or copies of published versions of their articles into such servers, thus providing readers worldwide with a free way of obtaining access to these papers, without needing paid subscription access to the source electronic journals."[12] Although supporters of these goals

7. The term *eprint* is here used to refer to any paper published in an electronic archive. In the literature, *preprint* usually refers to papers that have not yet been formally published elsewhere in a print or electronic journal, while *post-print* refers to electronically archived versions of articles that have been published elsewhere.
8. *See* Diann Rusch-Feja, *The Open Archives Initiative and the OAI Protocol for Metadata Harvesting: Rapidly Forming a New Tier in the Scholarly Communication Infrastructure*, 15 LEARNED PUBLISHING 179, 180 (2002).
9. *See* Paul Ginsparg, Creating a Global Knowledge Network, Address at the Second Joint ICSU Press-UNESCO Expert Conference on Electronic Publishing in Science (Feb. 20, 2001), *at* http://arxiv.org/blurb/pg01unesco.html; Paul Ginsparg, Winners and Losers in the Global Research Village, Address at UNESCO conference in Paris (Feb. 21, 1996), *at* http://arXiv.org/blurb/pg96unesco.html.
10. Most of Harnad's writings, as well as links to online discussions of the issues he raises, can be found at his Web site at the University of Southampton. *See* Harnad Papers on Online Research Communication and Open Access, *at* http://www.ecs.soton.ac.uk/~harnad/intpub.html (last visited July 5, 2004).
11. *See, e.g.*, Peter Suber, *Removing the Barriers to Research: An Introduction to Open Access for Librarians*, 64 C. & RES. LIBR. NEWS 92, 94 (2003). Suber also edits the monthly *SPARC Open Access Newsletter, at* http://www.earlham.edu/~peters/fos (last visited July 28, 2004).
12. Clifford Lynch, *Metadata Harvesting and the Open Archives Initiative*, ARL BIMONTHLY REP.. Aug. 2001, at 1,1, *available at* http://www.arl.org/newsltr/217/mhp.html; *see also* Rusch-Feja, *supra* note 8, at 180 (discussing "the philosophy of free access to information which has been promulgated by many proponents of full-text (e-print) servers").

    An "open archive" need not necessarily make its content available free of charge. It is possible for commercial publishers to incorporate open technical protocols to make their servers interoperable and their publications archives accessible through the same means as the free access servers. *See* Leslie Chan & Barbara Kirsop, *Open Archiving Opportunities for Developing Countries: Towards Equitable Distribution of Global Information*, ARIADNE, Dec. 2001, *at* http://www.ariadne.ac.uk/issue30/oai-chan/; Stephen Pinfield, *Open Archives and UK Institutions*, 9 D-LIB MAG. (Mar. 2003), *at* http://www.dlib.org/dlib/march03/pinfield/03pinfield.html.

often refer to eprint servers as "repositories" or "archives," they do not use the terms in the technical sense in which they are used by those concerned primarily with preservation of digital (or print) media.[13]

¶6 Ginsparg himself is recognized for establishing the best known and most successful eprint server (for researchers in high-energy physics and other scientific disciplines) at the U.S. National Laboratory in Los Alamos, New Mexico.[14] Ginsparg's server, known as arXiv.org (www.arXiv.org/), allows researchers to post or "publish" their research without utilizing traditional processes of peer review. To post a paper on the arXiv.org server, a researcher prepares his or her work in one of a number of accepted formats, then submits the paper by e-mail or FTP, or through forms on the arXiv.org Web pages. Interested readers find out about new papers through e-mail notification services or by searching the site. Once posted, papers can be read and commented on by readers; authors can respond to comments, rework papers, and post revised versions of their work. Eventually, many of the papers are submitted and selected for formal publication in the standard journals in these fields; frequently, the final edited version is itself posted on the arXiv.org site.[15]

¶7 Most of the initial enthusiasm for open access approaches to scholarly communication was expressed by individual researchers and educators, but international interest in the Ginsparg server and other open archiving projects has also fostered the development of several, sometimes loosely organized, initiatives and programs to explore the broader applicability of the approach. Following Ginsparg's own approach, these discussions were often framed in terms of discipline-level projects,[16] but more general standards-based initiatives have also been developed. The Open Archives Initiative (OAI) (www.openarchives.org) is an effort to develop interoperability standards for common interfaces and approaches, searching, and other services, with the end goal of facilitating international and interdisciplinary access to electronically published scholarship.[17] The group's Protocol for Metadata Harvesting[18] provides an appli-

13. *See* Lynch, *supra* note 12; Pinfield, *supra* note 12.
14. In 2001, Ginsparg and arXiv.org moved from Los Alamos to Cornell University, where the archive is part of Cornell's digital library initiatives.
15. *See* Stephen Pinfield, *How Do Physicists Use an E-Print Archive?* D-LIB MAG., Dec. 2001, *at* http://www.dlib.org/dlib/december01/pinfield/12pinfield.html; Cecelia Brown, *The E-volution of Preprints in the Scholarly Communication of Physicists and Astronomers,* 52 J. AM. SOC. INFO. SCI. & TECH.187 (2001). For the results of a study of chemists' (and chemistry journals editors') uses of a preprint server in chemistry, see Cecelia Brown, *The Role of Electronic Preprints in Chemical Communications: Analysis of Citation, Usage, and Acceptance in the Journal Literature,* 54 J. AM. SOC'Y INFO. SCI. & TECH. 362 (2003).
16. For discussion of a failed National Institutes of Health project to establish an electronic preprints server in biomedicine, see ROB KLING ET AL, THE REAL STAKES OF VIRTUAL PUBLISHING: THE TRANSFORMATION OF E-BIOMED INTO PUBMED CENTRAL (Oct. 3, 2001 [Apr. 4, 2002]) (CSI Working Paper No. WP-01-03, Version 4.41B-2), *at* http://www.slis.indiana.edu/CSI/WP/wp0103B.html.
17. For an introduction to the development and organization of the OAI, see generally Rusch-Feja, *supra* note 8.
18. The Open Archives Initiative Protocol for Metadata Harvesting (June 14, 2002), *available at* http://www.openarchives.org/OAI/openarchivesprotocol.html. The Mellon Foundation has provided $1.5 million in funding for seven institutions to create gateway or portal services using the OAI Metadata Harvesting Protocol. Donald J. Waters, *The Metadata Harvesting Initiative of the Mellon Foundation,* ARL BIMONTHLY REP., Aug. 2001, *available at* http://www.arl.org/newsltr/217/waters.html.

cation-independent framework for user communities engaged in publishing content on the Web. The evolving agreement on technical standards fostered a number of new initiatives. In 2001, a group of scientists organized under the name of Public Library of Science (PloS) called for a boycott of all journals that would not agree to publicly archive their articles within six months of the initial publication date, and by 2003 PloS was publishing its own peer-reviewed open access journals.[19] In February 2002, the Soros Foundation's Open Society Institute provided funding for the Budapest Open Access Initiative, which calls for free access to scholarly literature published in journals or in pre-print form, and recommends the development of OAI-compliant self-archiving tools for scholars and support for alternative journals committed to open access.[20]

¶8 Since late 2002, many of the ideas developed in the open access movement have gained greater currency and wider visibility through the emerging concept of *institutional repositories*, which was widely reported and discussed in the higher education and library communities and occasionally in the general media.[21] With the goal of including not only research papers, but course materials, data sets, and other forms of intellectual output as well, university-based institutional repositories would "gather as much of the intellectual output of an institution as possible in an easy-to-search online collection."[22] In November 2002, MIT and Hewlett-Packard announced the release of DSpace (http://dspace.org), MIT's open-source archiving

---

19. *See* PloS Open Letter, *at* http://www.plos.org/support/openletter.shtml (last visited Oct. 20, 2003). For background on PloS, see Helen J. Doyle, *The Public Library of Science: Open Access from the Ground Up*, 65 C. & RES. LIBR. NEWS 134 (2004). In October 2003, PLoS launched its first journal, *PLoS Biology*, with plans to begin a second journal, *PLoS Medicine*, in 2004. PloS journals are available in print and online. Costs of publication are borne by authors or their institutions rather than through subscriptions. *See* About the PloS Journals, *at* http://www.plos.org/journals/index.html (last visited July 28, 2004).

20. *See generally* Budapest Open Access Initiative, *at* http://www.soros.org/openaccess/read.shtml (last visited July 28, 2004).

21. *See, e.g.*, Jeffrey R. Young, *Superarchives Could Hold All Scholarly Output*, CHRON. HIGHER EDUC., July 5, 2002, at A29; Richard K. Johnson, *Institutional Repositories: Partnering with Faculty to Enhance Scholarly Communication*, 8 D-LIB MAG. (Nov. 2002), *at* http://www.dlib.org/dlib/november02/johnson/11johnson.html; Roy Tennant, *Institutional Repositories*, LIBR. J., Sept. 15, 2002, at 28; Clifford A. Lynch, *Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age*, ARL BIMONTHLY REP., Feb. 2003, at 1, *available at* http://www.arl.org/newsltr/226/ir.html; Deanna B. Marcum & Gerald George, *Can Electronic Scholarship Survive?* LIBR. ISSUES: BRIEFINGS FOR FAC. & ADMINISTRATORS, July 2003, at 1.

22. Young, *supra* note 21, at A29. In October 2002, ARL, SPARC, and CNI hosted a daylong conference on institutional repositories, featuring presentations on initiatives at the University of California, MIT, Ohio State, CalTech, and the University of Rochester, as well as remarks by Paul Ginsparg on the relationships between these efforts and his own project. The agenda for the conference and other materials from the conference can be seen at Institutional Repositories: A Workshop for Creating an Infrastructure for Faculty Library Partnerships, *at* http://www.arl.org/IR_agenda.html (last visited July 5, 2004).

SPARC has also issued two significant publications on institutional repositories: RAYM CROW, THE CASE FOR INSTITUTIONAL REPOSITORIES: A SPARC POSITION PAPER (2002); RAYM CROW, SPARC INSTITUTIONAL REPOSITORY CHECKLIST & RESOURCE GUIDE (2002); *see also* OPEN SOC'Y INST., A GUIDE TO INSTITUTIONAL REPOSITORY SOFTWARE (2003).

software.[23] In January 2003, six other universities joined MIT in forming the DSpace Federation to jointly test and improve the program with funding from the Mellon Foundation.[24] The Berkeley Electronic Press (BePress), which supports the University of California's eScholarship Repository (http://repositories.cdlib.org/escholarship/), supplies its digital repository software to institutional partners.[25]

¶9 In other developments stemming more directly from the activities of the Open Archives Initiative, successful experiments in establishing small-scale institutional eprint archives were reported using the University of Southampton's "eprints.org" software.[26] Some writers known for their skepticism about the general applicability of centralized discipline-wide approaches to eprint publishing are more enthusiastic about the possibilities of smaller scale repositories, supported by eprints.org and similar applications, in academic departments or schools.[27] By mid-2004, most of the questions for the institutional repositories movement seemed to involve the apparent reticence of faculty members to participate.[28]

Possibilities for New Forms of Scholarly Communication in Law

Electronic Journals in Law

¶10 As for most other materials of legal research in the United States, primary electronic access to scholarly literature is provided though LexisNexis and Westlaw. Although both services have developed substantial retrospective full-text libraries of law reviews and other periodicals, their contents consist almost exclusively of electronic versions of articles first published in print law journals.[29] There is little

23. *See* Sally Atwood, *MIT's Superarchive*, TECH. REV., Dec. 2002/Jan. 2003, *available at* http://www.technologyreview.com/articles/atwood1202.asp.
24. Dan Carnevale, *6 Institutions Will Help Fine-Tune a Popular New Archiving Program*, CHRON. HIGHER EDUC., Feb 14, 2003, at A36. For current information about DSpace Federation activities, see DSpace, Activities & Deliverables, *at* http://dspace.org/federation/activities.html (last visited July 5, 2004).
25. *See* Berkeley Electronic Press, BePress Repository Technology, *at* http://www.bepress.com/repositories.html (last visited July 5, 2004). BePress is an electronic publishing initiative headed by law faculty members at the University of California-Berkeley. As noted *infra* ¶ 17, BePress repository software supports the New England Law Libraries Consortium (NELLCO) Legal Scholarship Repository.
26. *See* William Nixon, *The Evolution of an Institutional E-Prints Archive at the University of Glasgow*, ARIADNE, no. 32 (June-July 2002), *at* http://www.ariadne.ac.uk/issue32/eprint-archives; Stephen Pinfield et al., *Setting Up an Institutional E-print Archive*, ARIADNE, no. 31 (Mar.-Apr. 2002), *at* http://www.ariadne.ac.uk/issue31/eprint-archives. For information about the eprints.org software, see EPrints, *at* http://www.eprints.org (last visited July 5, 2004).
27. *See* Rob Kling et al., *The Guild Model*, 8 J. ELECTRONIC PUBLISHING (Aug. 2002), *available at* http://www.press.umich.edu/jep/0801/kling.html.
28. *See* Andrea L. Foster, *Papers Wanted: Online Archives Run by Universities Struggle to Attract Material*, CHRON. HIGHER EDUC., June 25, 2004, at A37.
29. *See* WESTLAW DATABASE DIRECTORY 2004 EDITION 878–918 (2004); LEXISNEXIS 2004 DIRECTORY OF ONLINE SERVICES 239–45 (2004). For many journals, coverage in both services starts in the early or mid-1980s, although the Westlaw directory notes for a number of titles that coverage is selective until the early 1990s. For studies of uses of journal articles in the Westlaw and LexisNexis databases, see John P. Joergensen, *Second Tier Law Reviews, Lexis and Westlaw: A Pattern of Increasing Use*, LEGAL REFERENCE SERVICES Q., 2002, no. 2, at 43; Richard A. Leiter, *Use of Law Reviews in Modern Legal Research: The Computer Didn't Make Me Do It!* 90 LAW LIBR. J. 59 (1998).

evidence of movement toward either all-electronic journals or journals published simultaneously in both print and electronic formats.[30] At least one law school (Duke) regularly publishes its journals simultaneously on the Web and in print,[31] and at least one publisher of new electronic journals in law (BePress) has emerged.[32] Yet, it is probably no exaggeration to say that the major recent development in electronic publication of legal scholarship is Hein Online's retrospective database of law review articles.[33]

Working Papers Initiatives

¶11 The relatively sparse development of electronic journals in law might indicate that legal scholars have little interest in alternatives to the print journal form for disseminating scholarly legal information. Yet, legal scholars have demonstrated growing support for other forms of electronic publication. These include the postings of papers by individual authors on the Legal Scholarship Network (LSN) (http://www.ssrn.com/lsn/index.html)[34] and the increasing numbers of Web-based law school working papers series, published either directly on a law school's own Web site,[35] or through a third-party provider such as LSN or the Berkeley

30. Perhaps because of the central place that student-edited journals occupy in the publication of legal scholarship, there has been little development in law either of enhanced Web versions of existing journals or of new all-electronic journals. For fuller discussions of the impacts of electronic media on publication of legal scholarship, see Richard A. Danner, *Electronic Publication of Legal Scholarship: New Issues and New Models*, 52 J. LEGAL EDUC. 347 (2002); Marguerite Most, *Electronic Journals in the Academic Law Library—Law Reviews and Beyond*, LEGAL REFERENCE SERVICES Q., 2002, no. 4, at 189. *See also* CAROL TENOPIR, USE AND USERS OF ELECTRONIC LIBRARY RESOURCES: AN OVERVIEW AND ANALYSIS OF RECENT RESEARCH STUDIES 13 (2003) (reporting survey results suggesting that law professors rely less on electronic journals for research and teaching than scholars in other disciplines).

    For information on the number and characteristics of all-electronic journals in law, see Deanna Barmakian, Electronic Law Journals: The Invisible Literature, *at* http://www.law.harvard.edu/library/services/research/elj (materials prepared for presentation at AALL Annual Meeting, July 14, 2004) (last visited July 28, 2004).

31. Duke has published all six of its student-edited journals in Web versions since 1996. The Web site (http://www.law.duke.edu/journals/) provides a fairly straightforward replication of print versions of articles in HTML, with some (mostly internal) linking to sources cited and star-pagination to the print versions. Downloadable PDFs of the articles are also available. The all-electronic *Duke Law & Technology Review* (http://www.law.duke.edu/journals/dltr/) began publication in 2001.

32. BePress publishes several journals of legal scholarship: the three-part *Global Jurist* (http://www.bepress.com/gj/); *Issues in Legal Scholarship*, based on symposia focusing on key articles (http://www.bepress.com/ils/); and a "continuous" Web version of *Theoretical Inquiries in Law* (http://www.bepress.com/til/), which also appears in print. As noted earlier, BePress is also a provider of software for digital repositories.

33. For lists of its current and planned journal coverage, see HeinOnline, *at* http://www.heinonline.org (last visited July 5, 2004).

34. LSN is a division of the Social Science Research Network (SSRN), *at* http://www.ssrn.com (last visited July 5, 2004), which includes several other components (e.g., accounting and economics) in addition to law.

35. Bernard Hibbitts's Jurist Web site includes access to both pre- and post-print postings of articles on individual faculty members' home pages. *See* Jurist, Law Professors, *at* http://jurist.law.pitt.edu/home_pgs.htm (last visited July 5, 2004). Jurist also provides links to several models for working papers series published at law schools. *See* Jurist, Working Papers, *at* http://jurist.law.pitt.edu/ol_artcl.htm (last visited July 5, 2004).

Electronic Press (BePress) (http://law.bepress.com/repository). The growing amount of scholarship published electronically in working papers series reflects both the law schools' institutional interests in promoting the scholarly contributions of their faculties, and law professors' interests in more actively participating in the dissemination of their work. Like other scholars, law professors presumably "write for impact, not for money, [and] want the widest possible dissemination of their work."[36] As a result, it is not surprising that they would recognize the benefits of Web-based electronic publishing systems.

¶12 Proponents of open archiving are quick to point out that scholarly communications systems perform a number of roles: they provide quality certification, enable distribution and access, index the works, and provide means for archiving them. In an electronic publishing environment, these roles can be separated or "decoupled" in ways that they cannot be in an environment based in print packages. Most importantly, the quality certification that typically comes with having an article accepted for publication in a print (or electronic) journal can be separated from distribution and access, and need not be lost if authors choose to self-publish their papers on the Web or post their work on a law school server. There is no necessary reason to consider eprint posting and journal publication as exclusive alternatives, or for legal scholars interested in posting their finished work on an eprint server to forego publication in a leading law review. Indeed, the 1998 American Association of Law Schools "Model Author/Journal Agreement" allows authors to post works accepted for journal publication "on an Internet or Intranet site over which the Author has effective control."[37]

¶13 Law professors' interests in alternative means for disseminating their works are evidenced by the popularity of the Legal Scholarship Network, and by the recent entry of BePress into the market as a competing service. Established in 1997, LSN now publishes about fifty electronic subject matter "journals" under titles such as "Administrative Law," "Contracts and Commercial Law," or "Cyberspace Law," as well as working papers series published under the names of individual law schools.[38] LSN's approach and features are not unlike those of the arXiv.org server: papers (usually published articles or papers accepted for publication) are submitted for posting by faculty members or their schools, and e-mail "journals" alert subscribers to new papers and provide links to the full text of the papers on the LSN Web site. Institutional site licenses entitle law faculty members to subscribe to any number of the LSN e-mail journals; and the papers are also

36. Suber, *supra* note 11, at 94.

37. Attachment to Memorandum from Bari Burke, to Deans of Member and Fee-Paid Schools, Model Author/Journal Agreement (Memorandum 98-24, May 18, 1998), *available at* http://www.aals.org/9824.html.

38. LSN publishes more than sixty law school working papers series, some focusing on law and economics; most on public law and legal theory. A number of schools participate in both series. A Legal Studies Research Paper Series was established more recently. *See* Legal Scholarship Network Research Paper Series, *at* http://www.ssrn.com/lsn/index.html (last visited June 28, 2004).

searchable and downloadable in PDF format through the LSN Web site. Nearly all U.S. law schools have site licenses for LSN, as do a large number of law schools outside the United States and some law firms and corporations.[39] Papers stored on the LSN site are open to any user of the Web and can be searched by author, abstract/article title, or words within the text of the abstract. The BePress Legal Repository, established in 2004, includes listings of papers by subject area and by collection name. Searching can be done both by subject area and through the entire repository. Papers can be downloaded in PDF format, and e-mail notification services are provided. Although most of the papers available from both LSN and BePress have already been published in law journals, the BePress Legal Repository also includes a series of unpublished papers.[40]

¶14 LSN and BePress both serve to approximate for law the repositories and archives being developed in other fields. The popularity of LSN for authors of legal scholarship is evidenced in its tracking of the number of downloads for posted papers,[41] and BePress's recent entry into the market for legal scholarship as a competing service further supports the case for the applicability of the eprint server model to law.

Electronic Archiving of Legal Scholarship

¶15 Law professors, of course, do not have to rely on for-profit journals or other commercial services to publish and disseminate their scholarship. The primary vehicles for publishing legal scholarship continue to be the student-edited law reviews published at law schools. Yet, the rapid development of Web sites at nearly all U.S. law schools has lowered the barriers both for schools wishing to publish journals and working papers series electronically, and for individual faculty members wishing to self-publish their work in advance or independently of its publication in more traditional venues. Institutional support for electronic publishing initiatives and encouragement of experimentation can provide an effective counterweight to commercial development of the electronic publishing space for legal scholarship, keep prices low, and help law schools maintain and improve their traditional roles in publishing and distributing the literature of the field.[42]

---

39. A list of more than 370 LSN site license holders is posted on the SSRN Web site. Legal Scholarship Network Site Subscriptions, *at* http://www.ssrn.com/update/lsn/lsn_site-licenses.html (last visited June 28, 2004).
40. *See* ExpressO Preprint Series, *at* http://law.bepress.com/expresso/eps (last visited June 28, 2004).
41. As of July 5, 2004, LSN's most popular paper had been downloaded more than 35,000 times. *See* All Time Hits: SSRN Top Ten Downloads for Legal Scholarship Network, *at* http://papers.ssrn.com/toptens/tt_ntwk_201.html (last visited July 5, 2004). BePress also provides a list of popular papers. *See* The 10 Most Popular Articles in the ExpressO Preprint Series, *at* http://law.bepress.com/expresso/eps/topdownloads.html (last visited July 28, 2004).
42. The ready availability of the World Wide Web as a means of publication at law schools and universities has already created alternative publishing opportunities for law schools and for individual scholars. The most prominent examples are Cornell Law School's Legal Information Institute (http://www.law.cornell.edu/) and the Jurist site created by Bernard Hibbitts at the University of Pittsburgh School of Law (http://jurist.law.pitt.edu/).

¶16 The LEDA project (Legal Education Document Archive), a cooperative effort of the Harvard Law School Library, the Cornell Legal Information Institute, and several other law schools, is an example of an initial experiment in law.[43] LEDA was designed as a Web-based electronic publishing system for law schools to use on local eprint servers to hold and preserve whatever varieties of legal scholarship (working papers, published articles, student papers, theses, briefs) a school chose to include.[44] Like arXiv.org, the LEDA software allows authors to post documents directly (or with assistance) and make them readily available to others. In contrast to arXiv.org, however, LEDA envisioned not a centralized digital archive of legal scholarship, but a system based on distributed, but linked, Web servers housed at individual law schools. Searches can be performed either on a local server alone or simultaneously on all LEDA servers. In this way, LEDA recognized the traditional role of localized school-based publication in the discourse community of law.

¶17 Of course, there is also nothing to stop individual law professors from publishing their own scholarship and that of others on their own servers or those of their schools, and there is nothing to stop individual law schools (or consortia) from creating archives of scholarship based on infrastructure from suppliers such as BePress or eprints.org. The NELLCO Legal Scholarship Repository (http://lsr.nellco.org), which is supported by BePress technology, offers working papers from several law schools. Users can search either the entire repository or the contributions of individual schools. Duke Law School is using eprints.org technology for a project to provide persistent electronic access to the full body of scholarship produced by Duke faculty through a Faculty Scholarship Repository (http://repository.law.duke.edu). The eprints.org software is compliant with OAI standards and allows for deposit of papers in HTML, PDF, and other formats, as well as for links to materials published on remote servers.

Preservation of Born-digital Legal Scholarship

¶18 In the print environment, the forms of final publication of new scholarship provided the means to preserve the scholarship for future study and reference. The final, definitive version of a paper benefitted from review and acceptance before publication, as well as from editing, printing, packaging, and distribution in a journal, and "was, inevitably, the version referred to by subsequent authors."[45] In most disciplines, drafts and other informal means of communicating research results were not widely available for reading and comment.

---

43. *See* Harry S. Martin, *Legal Education Document Archive*, AALS NEWSL., Aug. 2002, at 8.
44. Harvard's LEDA server (http://leda.law.harvard.edu/leda) holds more than 500 papers, mostly by law students.
45. *Defining and Certifying Electronic Publication in Science: A Proposal to the International Association of STM Publishers*, 13 LEARNED PUBLISHING 251, 254 (2000) [hereinafter *Defining and Certifying Electronic Publication*].

¶19 In an electronic publishing environment, original drafts, revised drafts, and comments from others preceding the final version of a paper can all be widely disseminated and made publicly available without the value added by the processes of traditional formal publication. Because of their wide initial availability and possible persistence even after a "final" version of the paper is published, at least some of these forms must themselves be considered "published" in one sense or another.[46] As a result, the concerns of those involved with preservation of legal information must include not only the technical issues of how to preserve electronic documents but such issues as who will preserve them, what versions should be preserved, how different versions are distinguished, and how to ensure persistent access.

¶20 There has been increased interest in and support for preservation of electronic journals in disciplines other than law. The Mellon Foundation has funded digital archiving projects at a number of research libraries;[47] Stanford and Sun Microsystems have collaborated in the LOCKSS (Lots of Copies Keep Stuff Safe) (http://lockss.stanford.edu/) project, aimed at creating a cooperative decentralized archive system;[48] and JSTOR's scholarly journal archive (www.jstor.org/) provides the basis for a centralized electronic archive of journal literature.[49] While the preservation interests of journal publishers may not necessarily match the interests of librarians,[50] the growth in importance of electronic journals has at least prompted the possibility for discussions about publisher partnerships with libraries.[51]

¶21 Less attention has been paid to ensuring that electronically published scholarship outside the journal form will be preserved for the benefit of later generations of researchers and be easily retrievable. As noted earlier,[52] the term *archive* has been used frequently and loosely in discussions of electronic publishing, without much consideration of what will be needed to create and maintain a stable environment for ensuring that materials deposited in electronic archives and repositories will be available in the future. While the original open archives initiatives, such as Paul

---

46. "[I]n the electronic environment [the final version of an article] represents one point on a potential continuum of communication. Other points on that continuum (such as preprints) are becoming increasingly common currency, and there is unlimited potential to add to or even change electronic content after it has been made available." *Id.* at 251.
47. For a discussion of the projects, see Digital Library Fed'n, Summary of the Projects and Their Progress, *at* http://www.diglib.org/preserve/ejpreps.htm (last updated 2002); *see also* Dale Flecker, *Digital Archiving: What Is Involved?* EDUCAUSE REV., Jan./Feb. 2003, at 10.
48. *See* Victoria A. Reich, *Lots of Copies Keep Stuff Save as a Cooperative Archiving Solution for E-Journals*, ISSUES IN SCI. & TECH. LIBRARIANSHIP, Fall 2002, *at* http://www.istl.org/02-fall/article1.html.
49. *See* Guthrie, *supra* note 1, at 60; The Challenge of Digital Preservation and JSTOR's Electronic-Archiving Initiative, *at* http://www.jstor.org/about/earchive.html (last updated Jan. 22, 2004).
50. *See generally* Flecker, *supra* note 47.
51. For a publisher's comments on the need for archiving partnerships with libraries, see Dick Kaser, *The Future of Journals*, INFO. TODAY, Mar. 2003, at 1 (comments of Elsevier executive Pieter Bolman).
52. *See supra* text accompanying note 13.

Ginsparg's ArXiv.org server, were highly praised for disseminating new knowledge, there was little discussion of how they might provide permanent storage and persistent access to the papers submitted,[53] and some advocates argued that preservation concerns should not be a priority.[54] Librarians' greater involvement in DSpace and other more recent institutional repository projects suggests that more attention will be paid to preservation issues,[55] but the problems remain knotty from both technological and cost perspectives.[56]

¶22 In addition to technical problems, the digital environment also creates the likelihood that drafts and other early versions of scholarly works will be widely disseminated electronically and will persist after a final authoritative version is published. How easy will it be for an online reader to distinguish the version of an article published in the June 2004 issue of the *Duke Law Journal* from versions published by the author in her own law school's LSN working papers series in December 2003, posted in earlier draft form on a LEDA server in September, and distributed as an even earlier draft to a law professor electronic discussion list in May? What about an October 2004 version posted to the LEDA server incorporating new developments on the topic or responses to comments made on the *Law Journal* version? Who will decide which of these versions should be preserved (and how)? Who will provide the metadata to make the documents accessible and distinguishable?[57] Who will ensure that even final versions of electronic publications persist and can be found by later researchers?[58]

¶23 To begin to deal more effectively with these problems in law, as law professors and other legal scholars take more advantage of opportunities to publish their work outside the journal form, it may be helpful to rethink what it means for a document to be considered "published" in the born-digital environment. When there is more than final publication to consider, new definitions might be

53. Perhaps because of their interests in open electronic access to scientific literature and rapid dissemination of new research, proponents of these initiatives place great confidence in the benefits of author self-archiving. *See, e.g.*, Budapest Open Access Initiative, *at* http://www.soros.org/openaccess/ (last visited June 28, 2004): "The initiative has been signed by . . . a growing number of individuals and organizations from around the world who represent researchers, universities, laboratories, libraries, foundations, journals, publishers, learned societies, and kindred open-access initiatives. We invite the signatures, support, and participation of the entire world scientific and scholarly community."

54. For a listing of questions about the utility of worrying about preservation attributed to Stevan Harnad, see Stephen Pinfield & Hamish James, *The Digital Preservation of E-Prints*, 9 D-LIB MAG. (Sept. 2003), *at* http://www.dlib.org/dlib/september03/pinfield/09pinfield.html.

55. *See* William G. LeFurgy, *Levels of Service for Digital Repositories*, 8 D-LIB MAG. (May 2002), *at* http://www.dlib.org/dlib/may02/lefurgy/05lefurgy.html; Lynch, *supra* note 21. For a short, but thoughtful, commentary on the preservation questions raised by DSpace, see McKenzie Smith et al., *DSpace: An Open Source Dynamic Digital Repository*, 9 D-LIB MAG. (Jan. 2003), *at* http://www.dlib.org/dlib/january03/smith/01smith.html.

56. For a recent discussion of the practical technical, organizational, and managerial issues, see Pinfield & James, *supra* note 54.

57. *See* Roy Tennant, *Digital Libraries: Metadata's Bitter Harvest*, LIBR J., July 2004, at 32.

58. *See* E. Dana Neacsu, *Legal Scholarship and Digital Publishing: Has Anything Changed in the Way We Do Research?* LEGAL REFERENCE SERVICES Q., 2002, no. 2–3, at 105, for a discussion of disappearing electronic sources cited in law review articles.

needed.[59] In other disciplines, attempts at redefinition have yielded lists of essential characteristics of "published" documents, which might provide a starting point in law as well. Yet, an examination of these efforts immediately reveals the problems involved in defining *publication* once one considers the questions raised by informal communications and by documents created for the electronic environment. It is probably easier to define publication in terms of such characteristics as including the means to differentiate draft versions from "definitive publication," providing persistent accessibility for documents even if their Web location changes, certifying authenticity, and guaranteeing that the document will not be withdrawn from a personal or institutional server,[60] than it will be to ensure that they are implemented in a publishing environment that encourages self-publication and local experimentation.

¶24 It has been argued that because, to date, most eprint files contain only text and static images, they will be relatively easy to preserve, with the key challenges being those posed by hardware and software obsolescence.[61] Yet, as researchers come to expect more from electronic documents than mere desktop delivery of flat files to download and print, and as more writers take advantage of the tools available to them, features that might be seen now as frills—universal and persistent links to cited references, regular integration of media elements into text, and sophisticated internal search and navigational systems—will be considered essential elements of scholarly papers. The costs and means for preserving these forms of electronic literature remain largely unexamined.[62]

¶25 Although scholars are concerned about the continuing availability of electronic journals and other forms of scholarship,[63] it is likely that they themselves will be less interested in the preservation issues raised by a burgeoning electronic grey literature than they have been in making use of electronic publishing and networked communications for disseminating and publicizing their works in both formal and informal forms.[64] Who will take on the responsibility for preserving

---

59. The report of an International Working Group notes that:
> Publication is the hard currency of science. . . . Scientists need to know the status of the information they encounter, whether they need to refer to it, critique it, or build on it to advance their own work. The document also needs to persist, since in science identifying a clear context for later responses is essential to maintain the quality and integrity of subsequent discourse.

Defining and Certifying Electronic Publication, *supra* note 45, at 253.

60. *See* Kircz, *supra* note 4, at 267–71; *Defining and Certifying Electronic Publication*, *supra* note 45, at 253–58.

61. *See* Pinfield & James, *supra* note 54.

62. For initial discussion of these issues, see ABBY SMITH, NEW-MODEL SCHOLARSHIP: HOW WILL IT SURVIVE? (2003), *available at* http://www.clir.org/pubs/reports/pub114/pub114.pdf.

63. *See* Janet P. Parker & Mark Sandler, *What Do Faculty Want?* NETCONNECT, Winter 2003, at 26, 27.

64. *See* Robin Peek, *The Tangled Scholarly Publishing Mess: Now That We've Digitized Most of the Journals, What Comes Next?* INFO. TODAY, Apr. 2001, at 36, 36 ("The technical concerns—such as article linking and archiving multimedia—that are necessary to make this whole system really work aren't going to generate the same type of interest from the scholarly community that the first great wave of digitization was able to accomplish.").

materials published in local archives and repositories? JSTOR's Electronic-Archiving Initiative argues that, as a general matter, "because digital resources are neither owned nor held locally, neither the library nor its parent institution is naturally compelled to develop the infrastructure required to preserve or care for the electronic resource."[65] However, it would seem that for locally posted working papers and similar materials, preservation responsibilities are likely to remain local. In law, these responsibilities will reside with law schools and law libraries.

65. The Challenge of Digital Preservation and JSTOR's Electronic-Archiving Initiative, *supra* note 49.