The Open Access Availability of Library and Information Science Literature

Doug Way

To examine the open access availability of Library and Information Science (LIS) research, a study was conducted using Google Scholar to search for articles from 20 top LIS journals. The study examined whether Google Scholar was able to find any links to full text, if open access versions of the articles were available and where these articles were being hosted. The results showed that the archiving of articles is not a regular practice in the field; articles are not being deposited in institutional or subject repositories at a high rate; and, overall, the percentage of available open access articles in LIS was similar to the findings in previous studies. In addition, the study found that Google Scholar is an effective tool for finding known LIS articles.

Simply stated, a primary goal of the open access (OA) movement is to make scholarly literature freely available online.[1] And since the movement's inception, libraries have been near the forefront, advocating for OA, establishing institutional repositories, creating and supporting the development of OA journals, working with faculty and researchers to promote OA, and educating their constituencies about the problems in the scholarly communication system. In many ways, this involvement likely grew out of the principles of librarianship that align with the OA movement and at the same time out of necessity based on the crisis in scholarly communication that has led to increasing pressures on library budgets. These roles would lead one to assume that scholars and practitioners in the field of Library and Information Science (LIS) consistently make their scholarship freely available online by posting them in institutional and subject repositories. To determine whether this is indeed the case, the OA availability of articles from 20 LIS journals published in 2007 was examined.

Background

There are two main routes to provide OA to scholarly literature: OA journals and through placing articles published in traditional toll journals in OA archives or repositories.[2] This second practice is generally called self-archiving and refers to the practice of the author going through the process of making preprints and postprints of their work available online by submitting it to an archive or repository, or by posting these works on their own Web sites.[3]

Doug Way is Head of Collection Development at Grand Valley State University; e-mail: wayd@gvsu.edu.

A number of studies have examined the prevalence and practice of self-archiving. In a large study of faculty, Swan and Brown found that 49 percent of authors surveyed had self-archived at least one article in the preceding three years.[4] In a smaller study, Watson similarly found that 48 percent of scholars at Cranfield University had made some of their work freely available on the Internet.[5]

In examining where those faculty who do self-archive deposit their materials, studies have found faculty are more likely to place materials on personal Web sites than institutional or subject repositories.[6] Those studies that have examined participation in institutional or subject repositories have also generally found low participation rates. While Watson[7] found that 43 percent of faculty in her survey had deposited materials in their university's institutional repository, larger studies by Swan and Brown[8] and Lawal[9] found much lower participation rates.

Recently, the idea of self-archiving as a solution to the scholarly communication crisis has come into question. Studies examining self-archiving in institutional and subject repositories have found that, even when faculty work is placed in repositories, it is often not placed there by the faculty member, a process often referred to as mediated deposit.[10] The failure of self-archiving to become a regular practice among scholars is discussed in these studies and elsewhere in the literature.[11] As some have begun calling for broadening the definition of self-archiving to take into account the growth and importance mediated deposit is playing,[12] it is very possible that such terminology will fall away and the focus will instead be on archiving of literature, regardless of who placed it in a repository. In the literature there will also likely be a move toward examining the comprehensiveness of the OA availability of literature.

Two studies that do examine the comprehensiveness of OA literature looked at the availability of articles in the sciences. In one study, Wren examined the availability of articles that appeared in 13 scientific journals published over a ten-year span that were indexed in Medline.[13] He found a correlation between a journal's impact factor and its availability on nonjournal Web sites; it appeared that, as time passed, articles were more likely to be available online. In a second study, Matsubayashi and others examined a large sampling of biomedical articles published in 2005 to determine their OA availability using Google, Google Scholar, and OAIster.[14] Their study found that 27 percent of articles examined were available OA and that 70 percent of those articles were found on journal Web sites, while very few OA articles were found in institutional repositories or authors' Web sites.

While there has been much written about OA practices, there is surprisingly little examining the field of Library and Information Science's literature. King and others found that librarians tended to have a better understanding of issues related to the crisis in scholarly communication than faculty.[15] Still, in a study of the self-archiving practices of LIS scholars, Coleman and Roback found that only 55 percent reported self-archiving, and few self-archived in institutional repositories or the field's two main subject repositories, dLIST and E-LIS.[16] Instead, the majority placed their works on Web sites, network drives, or laptops. These results are in line with those found by Swan and Brown[17] and give credence to those who have criticized the lack of self-archiving in the LIS field.[18] In fact, Salo argues that "when asked to deposit their own content, librarians are as balky as faculty."[19] This is in spite of the fact that research has shown that the vast majority of the top LIS journals do not prohibit self-archiving[20] and that the majority of academic librarians believe the principles of the OA movement are related to the purpose of academic libraries.[21] Still, in spite of the criticism, there has not been a study that examines the OA availability of articles in the LIS literature.

Methodology

Following the methodology and rationale outlined by Coleman[22] in December 2008, a search was conducted in *Ulrich's Periodicals Directory* for those journals with a descriptor of Library and Information Science. From that list of journals, the twenty journals with the highest impact factor, as listed in the 2007 *Journal Citation Reports Social Science Edition*, were selected for evaluation.

Articles that appeared in journals' volume published in the year 2007 were selected for analysis. Articles for the individual journals were identified by visiting publisher Web sites where basic bibliographic information, including title and author for every article in all the journals, was found to be available. Research articles and essays were identified for inclusion in the project, while regular columns written by one individual, letters, errata, editorial comments, and so forth were excluded from the evaluation. A total of 922 articles were selected for inclusion in the study after reviewing each journal.

To determine the OA availability of these journal articles, a search was performed in Google Scholar for the title of the article. If an article had a common or simple title that returned significant numbers of results, additional limiters, primarily the author's name, were added to the search string to narrow the results. Search strategies were also modified when searches returned no results, often by either dropping a subtitle from the search string or by changing the search from "with the exact phrase" to "with all of the words" in the Google Scholar Advanced Search.

Even though it is possible to discover OA articles using a variety of resources, such as Google, CiteSeer^{X} or OAIster, Google Scholar was chosen as studies have shown that the vast majority of scholars regularly use it in their research. A study of academic science researchers at the University of North Carolina at Chapel Hill found researchers are increasingly using resources like Google Scholar to search across multiple resources.[22] A survey of faculty in eight disciplines across the sciences, social sciences, and humanities at Louisiana State University at Baton Rouge found that 90 percent of the respondents would use Google Scholar in their research.[24] A third survey of LIS scholars found that the majority used Google Scholar to search for OA content.[25] In addition, Markland found that Google Scholar did a better job finding items in institutional repositories than the traditional Google search.[26]

For each article, it was noted whether an OA version of the article was found, if only a citation for the article was found, or if no reference to the article was found. If an OA version of the article was found, it was noted whether the article was found on a faculty or scholar's Web site, an organizational Web site, in an institutional repository, in a subject repository, or on some other kind of Web site. For the purpose of this study, institutional repositories were defined as those organized by an institution, such as the University of Pennsylvania's ScholarlyCommons@Penn and the University of California's eScholarship Repository, while subject or disciplinary repositories were defined as those organized by a discipline, such as arXiv in Physics.[27] To examine the use of the two major LIS subject repositories, it was also noted if an article was available via dLIST or E-LIS. In the instance of one journal, it was determined that multiple volumes of that journal, along with those of other journals, had been systematically downloaded and posted on a freely available Web site that had been indexed by Google Scholar. Because those articles were posted in what readily appeared to be a violation of a license agreement, all the articles found on that Web site were excluded from the results analysis.

Results

Article Availability in Google Scholar

Links to full text were found for more than 86 percent of the 922 articles. This

TABLE 1
Location of OA Articles

	IR	SR	Personal Web Site	Organizational Web Site	Other	Multiple Locations
Percentage for All Articles (n = 922)	4%	10%	8%	3%	5%	3%
Percentage for OA Articles (n = 253)	14%	38%	29%	11%	18%	11%

number includes links to the full text on toll sites such as Elsevier's ScienceDirect in addition to OA versions of articles. Of the remaining articles, citations were found for more than 11 percent of the articles; and, overall, only 21 articles, or just over 2 percent, were not found in Google Scholar.

OA Availability

Of the 922 articles examined, OA versions were found for 253 articles. Table 1 shows the percentages of articles found in institutional repositories, subject repositories, Web sites maintained by authors, Web sites maintained by departments or organizations or related to a project and

TABLE 2
Percentage of OA Articles by Journal

Journal Name	Total # of Articles	% OA
Annual Review of Information Science and Technology	13	23%
ASLIB Proceedings	37	13%
College and Research Libraries	31	68%
Health Information Libraries Journal	38	0%
Information Processing & Management	108	30%
Information Society	33	18%
Interlending and Document Supply	34	12%
Journal of Academic Librarianship	68	12%
Journal of Documentation	41	24%
Journal of Information Science	48	17%
Journal of Information Technology	36	6%
Journal of the American Society for Information Science and Technology	187	31%
Journal of the Medical Library Association	64	94%
Law Library Journal	30	70%
Library and Information Science Research	25	12%
Library Quarterly	18	17%
Library Resources and Technical Services	20	10%
Online Information Review	42	5%
Portal	25	8%
Serials Review	24	13%

a category for articles found on Web sites that fall into other categories. Table 1 also shows the percentages of articles that were found on more than one site. In addition to those results, the study found that, overall, 1 percent of all articles were found in E-LIS and 2 percent were found in dLIST. Of the OA articles identified, only 5 percent were found in E-LIS and 7 percent in dLIST.

No journal, including the three OA journals included in this study, had OA versions for 100 percent of their articles. The percentage of OA articles found varied from zero to nearly 94 percent. Table 2 details the percentage of OA articles found for each journal examined.

Discussion

Lack of OA LIS Literature

An examination of the study results suggests that the field of Library and Information Science has failed to embrace archiving and OA as a regular practice. Articles are not being deposited in institutional and subject repositories at a high rate, and the fact that the vast majority of items are deposited in only one repository suggests that, among those who archive, there is little concern about preservation and redundancy. While the results suggest some disciplinary differences in the places where OA articles can be found, as well as in archiving practice, the overall percentage of OA articles is similar to that of previous studies.

Previous studies examining self-archiving found that nearly 50 percent of faculty reported doing some self-archiving, and surveys of LIS scholars found an even higher rate of self-archiving.[28] Yet this study found OA versions of only 27 percent of the articles examined. This discrepancy could be because those surveys did not examine whether faculty regularly or consistently archived, but instead looked at any instance within a given period of time. The findings of this study suggest the need to further examine this area.

The percentage of available OA articles is the same, though, as the findings of Matsubayashi and others in their study of the biomedical literature.[29] However, while their study found most articles were available on publisher Web sites, this study found that the most common place for OA works to be found was departmental or organization Web sites.[30] Only 36, or 14 percent of all OA articles, were found on publisher Web sites. This would suggest that archiving is more prevalent in LIS than in the biomedical science field, but it also suggests that literature in LIS is much less open to embargoed access.

Previous studies found that faculty were more likely to post their works on a personal Web site.[31] The results from this study would seem to contradict that; but it is important to keep in mind that those studies were focused on self-archiving behavior, whereas this study focused instead on availability. This study found subject repositories the most common place to find articles, with 96 articles found in them, compared to 74 articles found on personal Web sites. Of the 96 articles in subject repositories, 60 of them were from the *Journal of the Medical Library Association* whose articles are included in PubMedCentral, which was categorized as a subject repository. If those articles were excluded from the results, the results would more closely match previous studies, with more articles being posted on faculty Web sites than in institutional or subject repositories. At the same time, it is possible that those articles would have been posted on a personal Web site had they not be published in an OA journal. What is clear is that additional studies of OA availability are necessary to clarify this topic.

This study found few articles deposited in institutional repositories or the discipline's two major subject repositories, dLIST and E-LIS. There has been steady growth in the number of institutional repositories established over the past several years,[32] yet LIS scholars seem as unwilling as scholars in other disciplines to deposit their works in these repositories. If professionals in LIS are unwilling

to archive their works in repositories, it should not be surprising that repositories face difficulties in recruiting content. What makes this issue somewhat more difficult to examine is the interdisciplinary nature of LIS where one can find articles published by librarians, LIS faculty, computer scientists, and others. Research examining the characteristics of those who do and do not make OA versions of their work available will help scholars gain a better understanding of the issue.

Only 11 percent of the OA articles identified in this study were available or archived in more than one location. This is somewhat surprising considering the importance that librarians generally place on preservation. On the other hand, this may suggest and reinforce the idea that archiving and providing OA to articles is not done with a long-term view and that issues of preservation and long-term access do not enter into consideration. Yet, with most repositories in their infancy, there are sure to be some bumps along the road and inevitably some repositories will close just as some journals cease to publish. By making this content available in multiple venues, authors not only increase access to their works by providing more opportunities for a work to be indexed, but they also help to ensure that their works will be available online in the future.

Google Scholar as a Discovery Tool

The results of this study indicate that, at least in the field of LIS, researchers are likely to find a reference to an article using Google Scholar. Only 2 percent of articles searched returned no results. This would seem to substantiate Markland's conclusion that Google Scholar is an effective tool for finding articles when the journal article is known.[33] Still, of concern for OA advocates should be Google Scholar's lack of success in returning results for the three OA journals. The findings for the three OA journals examined can be found in table 3. While Google Scholar found an OA version of an article for 94 percent of the articles in the *Journal of the Medical Library Association*, those were all found in PubMedCentral and not on the journal's Web site. If this journal had not been included in PubMedCentral, then only those articles deposited in institutional repositories and on faculty Web sites would have been found in spite of the fact that this is an OA journal. Google Scholar did index articles found on the *College and Research Libraries* Web site and the *Law Library Journal* Web site, but it did not do a good job of capturing all of the journal content on those Web sites, leaving researchers with an incomplete picture of the available literature. Because this study included only three OA journals, it is difficult to say if this was an anomaly; but the issue warrants further consideration in future studies.

Conclusion

Providing access to information is a basic tenet of librarianship. Ranganathan's classic work, *The Five Laws of Library Science*, calls upon libraries to make information widely available and easily accessible to all people.[34] While Ranganthan's work referred to books, these principles hold true regardless of the format of the information and can

TABLE 3
Availability of OA Journals Via Google Scholar

Journal Name	OA Version Found	Citation Only in Google Scholar	Not Found via Google Scholar
College and Research Libraries	68%	29%	3%
Journal of the Medical Library Association	94%	2%	5%
Law Library Journal	70%	30%	0%

be seen in the field's support of the OA movement. Yet this study has found there is a seeming contradiction in the lack of archiving of articles appearing in the top LIS journals. This is in spite of the fact that a previous study found that 90 percent of these journals allow some form of self-archiving.[35] To remedy this, librarians and LIS scholars need to take a leadership role and set an example for other fields by archiving all articles they publish. The fact that library faculty are placing open access mandates on themselves is an encouraging development.[36] Even when not bound by a mandate, though, authors need to exert more control over their research through the use of addendums to publication agreements that allow them to retain certain rights, including the right to self-archive. If publishers do not allow self-archiving or accept addenda, authors then need to refuse to publish in those venues. It would be at that point that the availability of LIS literature would match the rhetoric of many of its practitioners, and the field could point to its own literature as a shining example of free and open access.

Notes

1. Raym Crow, *The Case for Institutional Repositories: A SPARC Position Paper* (Washington, D.C.: The Scholarly Publishing & Academic Resources Coalition, 2002), available online at www.arl.org/sparc/bm~doc/ir_final_release_102.pdf [accessed 17 December 2008]; Stevan Harnad, "The Self-archiving Initiative," *Nature* 410 (2001):1024–25; Stevan Harnad and Tim Brody, "Comparing the Impact of Open Access (OA) vs. Non-OA Articles in the Same Journals," *D-Lib Magazine* 10, no. 6; Stevan Harnad et al., "The Access/Impact Problem and the Green and Gold Roads to Open Access," *Serials Review* 30, no. 4 (2004): 310–14; John Willinsky, *The Access Principle: The Case for Open Access to Research and Scholarship* (Cambridge, Mass.: MIT Press, 2006); Jingfeng Xia and Li Sun, "Factors to Assess Self-archiving in Institutional Repositories," *Serials Review* 33, no. 2 (2007): 73–80.

2. Peter Suber, "Open Access Overview," June 2007. Available online at www.earlham.edu/~peters/fos/overview.htm. [Accessed 31 December 2008].

3. Crow, "The Case for Institutional Repositories."

4. Alma Swan and Sheridan Brown, *Open Access Self-archiving: An Author Study* (Cornwall, U.K.: Key Perspectives Limited, 2005). Available online at www.jisc.ac.uk/uploaded_documents/Open%20Access%20Self%20Archiving-an%20author%20study.pdf. [Accessed 15 December 2008].

5. Sarah Watson, "Authors' Attitudes to, and Awareness and Use of, a University Institutional Repository," *Serials* 20, no. 3 (2007): 225–30.

6. Anita Coleman and Joseph Roback, "Open Access Federation for Library and Information Science: dLIST and DL-Harvest," *D-Lib Magazine* 11, no. 12 (2005); Ibironke Lawal, "Scholarly Communication: The Use and Non-use of E-print Archives for the Dissemination of Scientific Information," *Issues in Science & Technology Librarianship* 36 (2002); Swan and Brown, "Open Access Self-archiving."

7. Watson, "Authors' Attitudes."

8. Swan and Brown, "Open Access Self-archiving."

9. Lawal, "Scholarly Communication."

10. Jengfeng Xia, "Assessment of Self-archiving in Institutional Repositories: Across Disciplines," *Journal of Academic Librarianship* 33, no. 6 (2007): 647–65; Jengfeng Xia, "A Comparison of Subject and Institutional Repositories in Self-archiving Practices," *Journal of Academic Librarianship* 34, no. 6 (2008): 489–95; Jengfeng Xia and Li Sun, "Assessment of Self-archiving in Institutional Repositories: Depositorship and Full-text Availability;" *Serials Review* 33, no. 1 (2007): 14–21.

11. Dorothea Salo, "Innkeeper at the Roach Motel," *Library Trends* 57, no. 2 (2008): 98–123; Mary Westell, "Institutional Repositories: Proposed Indicators of Success," *Library Hi Tech* 24, no. 2 (2006): 211–26.

12. Jengfeng Xia, "A Comparison of Subject and Institutional Repositories in Self-archiving Practices," *Journal of Academic Librarianship* 34, no. 6 (2008): 489–95.

13. Jonathan D. Wren, "Open Access and Openly Accessible: A Study of Scientific Publications Shared via the Internet," *BMJ: British Medical Journal* 330 (2005): 1128–31.

14. Mamiko Matsubayashi et al., "Status of Open Access in the Biomedical Field in 2005," *Journal of the Medical Library Association* 97, no. 1 (2009): 4–11.

15. C. Judson King et al., *Scholarly Communication: Academic Values and Sustainable Models*

(Berkeley, Calif.: The University of California, Berkeley, Center for Studies in Higher Education, 2006). Available online at http://repositories.cdlib.org/cshe/CSHE-16-06. [Accessed 17 December 2008].

16. Coleman and Roback, "Open Access Federation."
17. Swan and Brown, *Open Access Self-archiving*.
18. Salo, "Innkeeper at the Roach Motel."
19. Ibid., 105.
20. Anita Coleman, "Self-archiving and the Copyright Transfer Agreements of ISI-ranked Library and Information Science Journals," *Journal of the American Society for Information Science & Technology* 58, no. 2 (2007): 286–96.
21. Christy L. Palmer, Emily Dill, and Charlene Christie, "Where There's a Will There's a Way? Survey of Academic Librarian Attitudes about Open Access," *College and Research Libraries* 70, no. 4 (2009): 315–35.
22. Coleman, "Self-archiving."
23. Bradley M. Hemminger et al., "Information Seeking Behavior of Academic Scientists," *Journal of the American Society for Information Science & Technology* 58, no. 14 (2007): 2205–25.
24. Aaron Lercher, "A Survey of Attitudes about Digital Repositories among Faculty at Louisiana State University at Baton Rouge," *Journal of Academic Librarianship* 34, no. 5 (2008): 408–15.
25. Coleman and Roback, "Open Access Federation."
26. Margaret Markland, "Institutional Repositories in the UK: What Can the Google User Find There?" *Journal of Librarianship & Information Science* 38, no. 4 (2006): 221–28.
27. Suber, "Open Access Overview."
28. Swan and Brown, *Open Access Self-archiving*; Coleman and Roback, "Open Access Federation"; Watson, "Authors' Attitudes."
29. Matsubayashi et al., "Status of Open Access."
30. Ibid.
31. Swan and Brown, *Open Access Self-archiving*; Coleman and Roback, "Open Access Federation"; Lawal, "Scholarly Communication."
32. Karen Markey et al., *Census of Institutional Repositories in the United States: MIRACLE Project Research Findings* (Washington, D.C.: Council on Library and Information Resources, 2007), available online at www.clir.org/pubs/reports/pub140/pub140.pdf [accessed 10 January 2009]; Cat S. McDowell, "Evaluating Institutional Repository Deployment in American Academe since Early 2005: Repositories by the Numbers, Part 2," *D-Lib Magazine* 13, no. 9/10 (2007).
33. Markland, "Institutional Repositories in the UK."
34. Shiyali Ramamrita Ranganathan, *The Five Laws of Library Science* (Madras, India: The Madras Library Association / London: Edward Goldston, Ltd, 1931). Available online at http://dlist.sir.arizona.edu/1220/. [Accessed 28 July 2009].
35. Coleman, "Self-archiving."
36. Peter Suber, "The Library Faculty at Oregon State University Voted Unanimously to Adopt an OA Mandate," *SPARC Open Access Newsletter*, no. 132 (Apr. 2, 2009); Peter Suber, "An OA Mandate for U of Oregon Library Faculty," *Open Access News*, May 7, 2009, available online at www.earlham.edu/~peters/fos/2009/05/oa-mandate-for-u-of-oregon-library.html [accessed 17 May 2009]; Peter Suber, "An OA Mandate for Calgary's LCR Division," *Open Access News*, May 10, 2009, available online at www.earlham.edu/~peters/fos/2009/05/oa-mandate-for-calgary-lcr-division.html [accessed 17 May 2009]; Peter Suber, "An OA Pledge from the Gustavus Adolphus Library Faculty," *Open Access News*, May 17, 2009, available online at www.earlham.edu/~peters/fos/2009/05/oa-pledge-from-gustavus-adolphus.html [Accessed 17 May 2009].