Use of Web Resources in the Journal Literature 2001 and 2007: A Cross-Disciplinary Study

Li Zhang

This article examines Web resources in research articles from 30 scholarly journals in disciplines across the sciences, social sciences, and humanities. The purpose of the study is to report the degree to which scholars make use of Web-based resources in the journal literature and to identify Web citation characteristics within different subject areas. The study also explores whether any changes emerged between 2001 and 2007. The examination confirms the finding of previous studies that, even though Web resources are not heavily used in journal articles, the number of such resources is increasing. Publicly accessible database repositories and open source software prevail over other Web resources in research communication. The implications for academic libraries are discussed. The study suggests that new strategies need to be developed to manage Web-based information resources.

ince its advent in the early 1990s, the World Wide Web (Web) has transformed the way in which information is presented, disseminated, and used. The transformation has profoundly impacted the way people work and live. Without exception, the efficient and easily accessible features of the Web have also brought radical changes to the world of scholarship and impacted how scholars seek information, exchange thoughts and ideas, and share research findings. The increasing use of the Web as a medium for sharing information and the growing number of research materials published on the Web demand a greater understanding of the academic communication culture in the Web environment.

Review of Related Research

Shortly after the Web emerged, information scientists started to monitor the impact of electronic resources on the scholarly literature and the citation behavior of scholars during the process of research. In 1996, Stephen P. Harter and Hak Joon Kim studied citations in 279 articles published in peer-reviewed library and information science electronic journals (e-journals) to measure the extent to which authors cited e-journals and other online sources. Harter and Kim found that, while e-journals accounted for 0.2 percent of the references cited, citations of various online sources as a whole consisted of 1.9 percent of the total references.[1] In a similar citation study conducted in 2001, Yin Zhang also investigated the use of

Li Zhang is Assistant Professor/Science Librarian in the Mississippi State University Libraries; e-mail: lzhang@library.msstate.edu.

crl-98r1

Internet-based electronic resources by library and information science researchers for scholarly contributions. Zhang discovered that the percentage of electronic citations (e-citations) in print journals of library science increased from 0.2 to 5.2 percent during the period of 1991 and 1998, whereas the percentage of articles containing e-citations showed a dramatic increase from 1.8 to 33.9 percent.[2]

In 2002, Susan Davis Herring broadened the scope of disciplines for citation analysis. She examined electronic resources used in scholarly electronic journals that represented several interdisciplinary subject fields. In Herring's selected articles, which were published during a relatively short period of time (summer 1999 through spring 2000), more than 55 percent of the articles cited electronic references, which contributed to 16 percent of the total citations. Herring confirmed earlier studies that online resources had gained a wider recognition among researchers.[3] In the same year, by analyzing the Web resources cited in scholarly law review articles, Mary Rumsey reported that articles with Web citations increased substantially from 0.57 percent in 1995 to 23 percent in 2000. The number of Web citations per article also increased from 1.9 to 10.45 during those years.[4] Paul Wouters and Repke de Vries took a different approach, focusing on the role of hyperlinks (including references to the Web site of a journal itself) in formal scientific journals. They studied references in 38 journals in five scientific and social scientific disciplines for the years 1995, 1998, and 2000 and found that scholars and scientists did incorporate "hyperlinking references" into their research work. Their research results drew a conclusion that not only authors but also publishers could impact the use of hyperlinks connecting to other content in research publications.[5]

Lyn Robinson examined the impact of new forms of information resources, such as Weblogs and wikis, in the literature of toxicology in the years 2000 and 2005. Robinson's findings showed that the percentage of citations to Web documents increased noticeably between 2000 and 2005; however, new types of Web-based social communication resources did not play any role in the formal research discourse.[6] Cecelia Brown's study of the electronic preprints (e-prints) being cited and accepted in the literature of physics and astronomy revealed that e-prints were becoming increasingly important for physicists and astronomers in the cause of scientific research. Brown proposed that "scholars in other disciplines will likely follow suit and accept e-prints as they realize their enhanced utility and ease of use."[7] On the other hand, Brown's recent study of journals published by American Chemical Society found that chemists were not fully using Web-based information resources in their formal scholarly literature, although the number of Web-based references had shown growing increase over the past decade.[8]

Objectives and Research Questions of the Study

Even though some previous work (for example, the above-mentioned studies from Herring and Wouters and Vries) conducted comparative studies of citations to Web-based information resources among disciplines, there are still relatively few articles that have undertaken cross-disciplinary investigations, particularly in relation to the humanities, regarding the formal use of Web resources to support research activity. In the current paper, the term "Web resources" indicates the sources that are searchable on the Internet, including electronic journal articles, electronic books, databases, publications produced by government and institutions, conference proceedings, technical reports, pre-prints, personal homepages, course materials, e-mails, online discussion forums, wikis, blog postings, RSS feeds, podcasts, and other types of information. The purpose of the paper is to add to the body of knowledge about the impact of the Web on scholars' information use

and citation practices. By examining and comparing references to Web resources, the study attempts to provide insights into the characteristics and differences of Web citation practices in a range of disciplines. To identify disciplinary characteristics, the study focuses on six subject fields: chemistry, biology, economics, anthropology, history, and linguistics. Each of these fields ranked relatively high on the list that Henk F. Moed described relating to the ISI coverage of the written research literature in a field.[9] Meanwhile, these subjects represent a diversity of disciplines in the sciences, social sciences, and humanities.

It is widely acknowledged that differences exist in the nature and research methodologies of diverse academic disciplines. It is also generally recognized that the outlet of formal scholarly communication differs between the disciplines of the sciences, social sciences, and humanities. While scientific researchers favor journal article publishing for sharing research discovery in a timely manner, social sciences and humanities scholars tend to employ books as a better channel for disseminating intellectual output.[10] In the electronic environment, with the ever-growing number of Web information resources available in greater depth, a logical question follows: Do Web resources influence the citation behavior of scholars in different fields? This study will address the following three specific questions:

1. To what degree have Web-based resources been accepted and used in the journal literature? Are there disciplinary differences in incorporating these resources?
2. What are the characteristics of the Web resources cited across disciplines?
3. Are there recognizable trends in the use of Web resources?

Answers to these questions will help to reveal a relatively thorough picture of the Web citation habits of scholars, provide useful information to Web page producers and information science professionals, and serve as a basis for further bibliometric analysis of citations to and from Web resources.

Methodology

The data collected for this study came from articles published in top-ranked journals that were identified by the Institute for Scientific Information's 2007 Journal Citation Reports (the most recent edition available when the present study was begun in May 2008). In addition to high prestige and wide readership, these journals also cover a good range of subject aspects of each discipline.[11]

The study drew upon full-length peer-reviewed research articles in all issues of the 30 journals published in sample years 2001 and 2007. Excluded items for analysis were conference reports, book reviews, review articles, technical notes, feature articles, communications, editorial materials, news items, and short commentaries and replies. The study examined Web resources cited in the reference list as well as those mentioned in other locations within an article, such as in the text, figures, table footnotes, and appendixes. After lengthy manual verification, total number of the research articles published during the two years, number of articles with Web resources, number of total citations, and number of citations to Web resources, were recorded into two separate Excel files. The URL of each Web resource was also noted for determining the nature of the Web resource and its availability in 2009.

It is necessary to explain the classification methods used for classifying Web resources. The category "journals" included both subscription-based and open-access journals. References to statistical data, data series, and digital collections of historical records were placed in the category of "statistical datasets or archives," which appeared mainly in the social sciences and humanities literature. Large collections of protein *sequences*, structures, spectra, genomes, and research literature were categorized as "repository databases," which was the principal type of database in the science literature. While "technical

reports" were defined as research reports produced by organizations, institutions, and companies, "other grey literature" was used to describe the resources indicated as technical standards, working papers, posters, fact sheets, newsletters, trade literature, and whitepapers. "Research homepages" consisted of laboratory or research group homepages, academic department homepages, and personal (researcher/professor) homepages. "Educational materials" were generic documents created by institutions, research centers or laboratories, professional associations, and companies, providing information on research, learning, technology, laboratory procedures, use of online tools, and program instructions. Nomenclature was also grouped in this category. "Course materials," however, indicated more specific teaching documents, including syllabi and lecture notes. "Search engines" specified the Web sites that collect various contents and enable users to retrieve data from a range of databases. An example is Entrez, The Life Sciences Search Engine (www.ncbi.nlm.nih.gov/sites/gquery). "Web portals" referred to Web sites that present an array of resources, tools, and services. An example is ExPASy Proteomics Server (http://ca.expasy.org/). "Web 2.0 communication forms" indicated social applications, such as wikis, blogs, podcasts, online discussion forums, and mailing lists. The "other" category covered announcements, mission statements, depository request forms, product brochures, surveys, directories, interviews, organization Web sites, and company Web sites.

Results

Use of Web Resources in Academic Scholarship

The study investigated a total of 13,859 research articles published in the aforementioned 30 journals in six different disciplines of the sciences, social sciences, and humanities for the years 2001 and 2007. All six of the disciplines used Web resources in their peer-reviewed journal articles. Collectively, there were 2,060 articles containing at least one Web resource; thus, 14.9 percent of the total number of articles contained Web resources. The study further revealed that, in general, the research articles incorporating Web resources constituted approximately 10.1 percent of the total number of articles published in 2001 and 19.0 percent in 2007. Compared with frequently referenced traditional sources such as books and journals, the occurrence of Web resources in the scholarly work seemed insignificant; however, there appeared a greater growth in the use of Web-located resources across all six disciplines from 2001 to 2007, especially in the humanities field (see table 1). The finding of increase corroborates previous studies of using electronic resources in a variety of academic disciplines. A further detailed analysis uncovered that, in the individual subject fields of the study, the percentage of increase ranged from 6.9 percent in chemistry to 45.2 percent in history. In addition, the total 13,859 articles accumulated 633,069 citations for the sample years; and the average percentage of Web resources in the articles that included any of such references increased from 0.4 percent to 0.8 percent, with greater rates

| TABLE 1						
Changes of Proportion of Articles with Web Resources by Discipline (%)						
Year	**Sciences**		**Social Sciences**		**Humanities**	
	Chemistry	**Biology**	**Economics**	**Anthropology**	**History**	**Linguistics**
2001	6.7	20.9	18.9	5.9	12.9	6.1
Average	10.1		12.1		7.3	
2007	13.6	31.9	31.1	25.4	58.1	21.5
Average	17.7		28.0		30.6	

observed in social sciences and humanities journal articles. It should be pointed out that this work only considered those resources with Uniform Resource Locators (URLs) to be Web citations. In many of the articles examined, particularly in the chemistry articles, authors listed unpublished results, unpublished data, patents, or personal communication as references in the endnote section. It is difficult to verify whether these references were from print or electronic resources. It may be the case that the unpublished results or data came from working papers available on a researcher's Web homepage. Patents are likely to be consulted via Web sites such as the United States Patent and Trademark Office site. Personal communication may take place via online discussion forums or through e-mails. In the same vein, for citations in journals from the other five disciplines, the study was unable to take into account the number of resources that scholars may have read electronically but chosen to cite them in print format. If the information consulted via electronic modes were documented as Web-based resources, the proportion of Web references would have been larger. The factor that reading patterns and citation patterns may differ presents the importance for journal editors to develop a complete set of guidelines on citing resources found on the Web.

Continuity of Web Resources Content
The percentage of accessible Web resource content was relatively low. Among the 3,769 Web resources found in the journal articles, only about 80.9 percent (n = 3,049) of the referenced content was available at the URL cited in for the Web resources. However, through the Internet Archive, a "non-profit that was founded to build an Internet library…offering permanent access for researchers, historians, scholars, people with disabilities, and the general public to historical collections that exist in digital format,"[12] about 310 out of the 720 nonworkable Web URLs led to the content.

In general, the accessibility and stability of Web resources were related to the age of the articles. While Web resources appearing in earlier year (2001) of the publications displayed a higher proportion of broken links (36.2%), the percentage of invalid URLs was only 10.5 percent for the 2007 articles. Figure 1 presents the continuity of Web resources by subject discipline. In five disciplines—chemistry, biology, economics, history, and linguistics—the percentage of accessible Web content was higher for more recent articles, with the sciences literature showing the largest increase in accessibility from 61.4 percent in 2001 to 89.0 percent in 2007. Humanities literature followed, ranging from 70.1 percent in 2001 to 83.3 percent in 2007. The percentages of workable URLs in economics and anthropology journal articles did not show much difference between 2001 and 2007. In both years, more than 75 percent of valid Web resources were observed.

FIGURE 1
Continuity of Web Resources by Subject Discipline
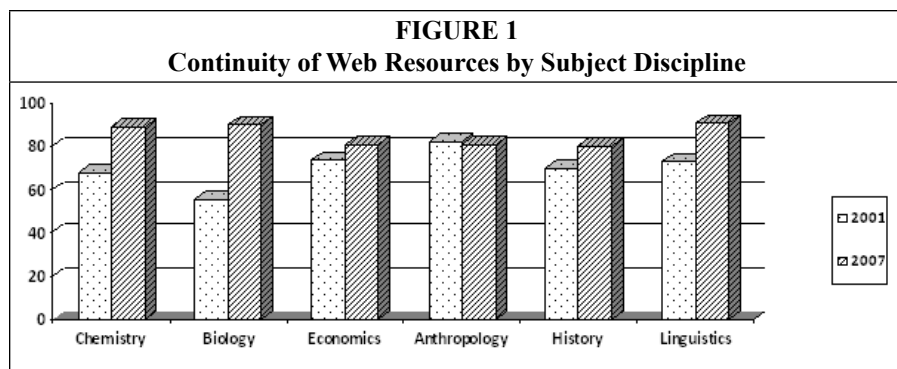
Characteristics of Web Resources across Disciplines

With the Web continuing to play an important role in research practices, analyzing the types of Web resources used in research articles can show the degree to which different types of Web publications impact academic literature. The present study classified all the active Web resources (n = 3,049) and found that scholars used a broad range of Web publications. Tables 2, 3, and 4 show the results of analysis of the Web resources in terms of the variety of source types.

In general, among all articles from the 30 journals, "conventional" resources on the Web appeared in the form of journals, books, government publications, technical reports, conference proceedings, working papers, theses or dissertations, newspapers, and magazines. "Unconventional" but noteworthy Web resources included publicly accessible repository databases, statistical datasets, software or Web-based computational tools, and personal or research group homepages.

Further examination of the Web resources from the chemistry and biology literature revealed that repository databases, Web-based analysis or computational programs, and open source software were the most heavily referenced Web resources. These references, largely produced by institutions and government agencies, made up about 50 percent of the cited Web resources, with 50.3 percent discovered in the 2001 samples and 53.8 percent in the 2007 samples. As noted in citation studies that analyze cited print resources, this study found that, when citing Web-based resources, scientists still tended to consult government publications (6.4%), technical reports (4.2%), and various other grey literature materials (7.7%), which mostly comprised technical standards and whitepapers. Among the remaining Web resources that carried a high percentage of references were research homepages (5.1%), educational materials (3.5%), and journal articles (3.1%). Homepages of researchers or research groups or academic departments usually provided additional relevant information about the researchers and ongoing research projects. Educational materials covered a wide variety of resources, including laboratory protocols, description of equipment, and guidelines for the use of online tools. When examining the overall citation content of the science articles, the study found that, as would be expected, journals (including both print and electronic forms) accounted for the predominant proportion of the total citations. However, electronic journals only made up a small percentage compared to its print counterpart in the citations. In light of the influx of social media sites, the study paid special attention to the degree to which Web 2.0 communication forms impacted scholarly research. A total of twelve citations (0.5%) were found coming from Web 2.0 communication forms, with seven being wikis and five being online forums and mailing lists. The rest of the Web resources on the whole covered a small percentage of the references.

As is seen in table 3, Web-based dataset repositories and analysis programs, which are the major forms of Web references in scientific research, are also important in economic and anthropological research. In this study, these two categories constituted 24.5 percent of the total citations to Web resources in social science journals. Government publications made up 13.5 percent of the cited Web resources; reports from organizations and companies, 7.4 percent; and other grey literature, 12.6 percent. It is noteworthy that working papers appeared to be a frequently used source; this is particularly the case for economists. In the 2007 data, approximately 43 percent of the Web resources from the "other grey literature" category were working papers. As was observed in the science literature, providing Web links to direct readers to personal or academic department homepages was also common for social sciences scholars for sharing their research with a wider intellectual community. The current study discovered 20 (about 15%) of such Web resources. In the social sciences lit-

TABLE 2				
Distribution of Types of Web Resources Cited – Selected Sciences Journal Literature				
	2001		2007		
Type	No.	%	No.	%	Mean (%)
Journals	11	2.2	65	3.4	3.1
Books/Book Chapters	0		15	0.8	0.6
Government Publications/Websites	25	5.0	131	6.8	6.4
Repository Databases	133	26.5	389	20.2	21.5
Analysis Tools & Software	119	23.8	648	33.6	31.6
Dissertations/Theses	1	0.2	5	0.3	0.2
Technical Reports	7	1.4	94	4.9	4.2
Conference Proceedings/Abstracts	3	0.6	9	0.5	0.5
Other Grey Literature	31	6.2	156	8.1	7.7
Research Home Pages	76	15.2	48	2.5	5.1
Educational Materials/Websites	34	6.8	52	2.7	3.5
Course Materials	4	0.8	3	0.2	0.3
Search Engines	3	0.6	15	0.8	0.7
Web Portals	5	1.0	27	1.4	1.3
Images	7	1.4	33	1.7	1.6
Web 2.0 Communication Forms	2	0.4	10	0.5	0.5
Video/Audio	1	0.2	1	0.1	0.1
Other	39	7.8	226	11.7	10.9
Total	501	100.1*	1927	100.2*	99.8*

erature, educational materials (7.1%) were more likely to be articles that provided background information, program instructions, additional references, and research aids. Although books and journal articles are conventional communication vehicles in social sciences research, the scholars did not fully incorporate the electronic formats of these sources into their research work (4.6% vs. 1.5%). The study did not explore whether this was due to insufficiency of such electronic materials for use, or to the information-seeking habits of the scholars (that is, their preference for print over electronic formats). Only 0.6 percent of the references to Web resources were Web 2.0 types of content, with those being to wikis and online discussion messages.

The reference patterns observed in the sample humanities literature is a little different. The most commonly cited Web resources were collection databases or archives, followed by newspapers, magazines, and news sites. The collections included databases of statistical information, pamphlets, fact sheets, audio clips, and various institutional digital archives or Internet libraries. Table 4 shows that a large percentage of references are to Web-based newspapers and magazines. The high use of newspapers and magazines may result from some specific research topics. In the history journals, a number of articles discussed Hurricane Katrina with many news sites included; the topic may have increased the number of references to such types of sources. About 8 percent of the references pointed to project reports, mostly produced by institutions and organizations, and the reports category

comprised the third largest Web reference resources. While electronic books only made up 2 percent of the total resources, journal articles represented higher portion of coverage, approximately the same as that of reports (7.5%). The fact that journal articles were often consulted in the history and linguistics literature corresponds with the finding of a citation study conducted by Helen Georgas and John Cullars on the linguistics literature. In their article, the authors claimed that, due to the prominent role that journals played, "the publication and citation patterns of linguistics more closely resemble those of the social sciences than the humanities."[13] Other prevalent sources that humanities scholars consulted were educational materials and government publications, constituting 6.4 percent and 6.1 percent of the Web resources, respectively. Although a very few Web 2.0 resources were found in the science and social science journal articles, about 2.0 percent of the Web resources in humanities articles were Web 2.0 oriented, with examples being blogs, wikis, and photo-sharing sites.

Discussion

The current study analyzed 3,769 Web resources in research articles from 30 peer-reviewed journals in disciplines spanning the sciences, social sciences, and humanities. Employing different sample sources

TABLE 3 Distribution of Types of Web Resources Cited – Selected Social Sciences Journal Literature				
	2001		**2007**	
Type	**No.**	**%**	**No.**	**%**	**Mean (%)**
Journals	3	4.2	12	4.7	4.6
Books/Book Chapters	0		5	2.0	1.5
Government Publications/Websites	13	18.3	31	12.2	13.5
Newspapers & Magazines	3	4.2	11	4.3	4.3
Statistical Datasets	8	11.3	37	14.4	13.8
Analysis Tools & Software	4	5.6	31	12.2	10.7
Dissertations/Theses	0		1	0.4	0.3
Reports	2	2.8	22	8.6	7.4
Conference Proceedings/Abstracts	0		2	0.8	0.6
Other Grey Literature	7	9.9	34	13.3	12.6
Research Home Pages	24	33.8	25	9.8	15.0
Educational Materials/Websites	2	2.8	21	8.2	7.1
Course Materials	0		3	1.2	0.9
Search Engines	0		0		
Web Portals	0		0		
Images	0		0		
Web 2.0 Communication Forms	0		2	0.8	0.6
Video/Audio	0		0		
Maps/Atlases	0		2	0.8	0.6
Other	5	7.0	16	6.3	6.4
Total	71	99.9*	255	99.9*	99.9*

TABLE 4				
Distribution of Types of Web Resources Cited – Selected Humanities Journal Literature				
	2001		2007	
Type	**No.**	**%**	**No.**	**%**	**Mean (%)**
Journals	6	6.2	16	8.0	7.5
Books/Book Chapters	2	2.1	4	2.0	2.0
Government Publications/Websites	7	7.3	11	5.5	6.1
Newspapers & Magazines	19	19.8	41	20.6	20.3
Datasets/Archives	20	20.8	44	22.1	21.7
Analysis Tools & Software	2	2.1	8	4.0	3.4
Dissertations/Theses	0		2	1.0	0.7
Reports	9	9.4	14	7.0	7.8
Conference Proceedings/Abstracts	1	1.0	2	1.0	1.0
Other Grey Literature	4	4.2	9	4.5	4.4
Research Home Pages	3	3.1	5	2.5	2.7
Educational Materials/Websites	8	8.3	11	5.5	6.4
Course Materials	0		0		
Search Engines	1	1.0	1	0.5	0.7
Web Portals	0		0		
Images	1	1.0	0		0.3
Web 2.0 Communication Forms	1	1.0	5	2.5	2.0
Video/Audio	3	3.1	3	1.5	2.0
Maps/Atlases	0		1	0.5	0.3
Other	9	9.4	22	11.1	10.5
Total	96	99.8*	199	99.8*	99.8*

and investigation approaches, this work obtains similar results to other studies and makes some unique discoveries as well.

The findings concur with the findings of other researchers that the number of Web resources in journal literature has grown over time, even though scholars appear to make limited use of Web publications. The increased number of Web resources in journals across all the six disciplines is an expected result. Due to the advancement of the Web, which has significantly expanded and improved the availability of research materials, the scholarly community is able to have a greater body of publications to reference in academic work. Yet, on the other hand, the inconsiderable propor-

tion of Web resources in these journals indicates that academics have not fully accepted the readily available Web resources and therefore incorporated them into their formal research. Possible reasons explaining the sporadic scatter of cited Web resources include the following:

(1) scholars do not perceive that Web documents are of scholarly value as peer-reviewed publications are;

(2) the citation patterns of scholars contrast with their information-seeking practices; they may have searched, read, and used online resources, but they either choose to cite them in their print copies if available or determine not to include every resource consulted in their work;

(3) there is a lack of established guidance in journals on citing various nontraditional types of materials, including Web resources; or

(4) scholars are uncertain or unaware of the criteria for evaluating Web resources, which prevents them from referring to reliable Web publications.

Identifying concerns and obstacles of using Web-based resources will help information professionals make plans and find appropriate solutions. For example, reference and instructional services librarians will thus need to help scholars become aware of scholarly Web resources, particularly informing them that some Web resources are peer reviewed. In addition to disseminating such knowledge at information desks, in classrooms, or at outreach places, librarians should develop online tutorials, subject guides, or video presentations showing the value of authoritative Web resources, instructing the criteria for evaluating Web resources and providing examples for citing Web resources. To reach a broader audience, libraries should strategically publicize the tutorials through multiple media channels. As more journals become openly accessible, librarians should carefully select quality journals, and cataloging librarians are required to effectively classify them. Additionally, user surveys could help libraries explore strategies for building reliable Web-based repositories and making them more visible and widely used.

The study's closer investigation of the dispersion and use rates of Web resources in journals from the six disciplines adds valuable data to similar research. It was predicted that scientists in chemistry and biology would be more likely to take advantage of the Web's fast and efficient characteristics in disseminating information than scholars in the social sciences and humanities. The results of this study, however, were contrary to the prediction. The average percentages of the Web resources in the chemistry and biology literature in both years were found to be lower than those in economics, anthropol-

ogy, history, and linguistics. In addition, pertaining to the articles containing Web resources, the lowest proportion appeared in the science journals in 2007, with the highest proportion observed in the humanities journals. The case in 2001 was different, with the science literature ranking second place in the average percentage of articles referring to Web resources. The disciplinary differences in the use of Web resources may express the degree of acceptance among scientists and scholars concerning the Web as a scholarly resource, or the differences may merely be evidence of scholars' information use habits. Also, it is worth mentioning that research focuses and designs may cause the varied distribution of Web resources in the six fields. It would be interesting to include more disciplines for analysis and find out the actual use patterns.

In contrast to the assumption that Web resources in research papers would remain more consistent, the study found that the overall valid ratio of active Web resources across the disciplines was 80 percent, ranging from a low of 55 percent observed in biology articles published in 2001 to a high of 90 percent in linguistics articles published in 2007. Not surprisingly, the discontinuity of Web resources is due to the dynamic nature of Web pages that leads to changed content or missing information. According to Diomidis Spinellis, the half-life of a referenced URL is about four years since the date it was published.[14] The poor stability and persistence of Web references may be one of the factors that discourage scholars in using the resources available on the Internet for scientific publication.

As the Web continues to be an essential research and communication tool, the decay of Web resources in scientific research presents a critical issue to resolve. It demands efforts from authors, journal editorial boards, and publishers to create and maintain an active online archive of Web resources. The Internet archive, Way-BackMachine (www.archive.org/web/web.php),[15] is an example that offers permanent

access to digital collections of research materials. To achieve the goal of preserving cited Web content in research articles, journal editors need to develop complete and specific guidelines instructing authors to document Web references. Authors should provide accurate and up-to-date bibliographic information in citing Web resources. If possible, they might attach printed copies of the cited Web pages to their submission package to ensure that future readers will be able to access the information. (This way, however, authors need to request copyright permission from the Web page creator to use printed copies.) Publishers should work with authors and editorial boards to store Web-based references in a stable depository that scholars and readers can draw on.

In terms of the types of Web resources employed in all sample journal articles, publicly accessible repository databases or archives, open source analysis software, and research homepages were the most frequently referenced materials. Newly emerged communication vehicles, including wikis, blogs, and mailing lists, did not receive a high rate of citation. This implies that scholars have concerns about the scholarly impact of these new forms of Web content. Other types of Web resources, such as electronic journals, books, and dissertations/theses, which are considered to be conventional forms of information research sources, accounted for only a small portion of the total references. Given the increasingly significant number of digitized journals, monographs, dissertations and theses available, why do scholars tend to make minimum use of electronic resources? Do they benefit from libraries' initiatives of electronic delivery of information? Library use data generally shows much higher use for electronic resources in the library's collection than print resources. Why did citation patterns not reflect this? The observations gained from the study present inquiries for further exploration. Answers to these questions may help academic librarians understand the concrete information needs of their user community, stimulate them to investigate new approaches for promoting electronic resources, and aid them when examining print collection policies and making decisions.

As is evident in the literature, open access has been one of the research focuses of information science studies. The data collected in the study might contribute to the efforts aimed at investigating open-access models. The study found that the prevailing number of Web references was not library-licensed but openly available to the public. These kinds of resources occur in the form of long-standing types of grey literature, such as conference proceedings, various reports, and working papers. These resources also occur in the form of newer content platforms, which are open-access repositories. There were three major types of content in the open repositories: data collections, peer-reviewed journals, and computational software. As is mentioned earlier in the study, repository datasets and open source software were heavily consulted open-access resources. Furthermore, the proportion of references to these two types of repositories showed a compelling increase from 2001 to 2007. On the other hand, among the three distinct types of repositories, the study did not find a significant number of references to open-access journals across the disciplines, nor did it observe a clear trend indicating that references to open-access journals increased substantially from 2001 to 2007; indeed, only the science literature showed a slight growth (from n = 5 to n = 21), with negligible growth in the social sciences and humanities literature (from n = 0 to n = 4, n = 4 to n = 6, respectively). For such references, authors referred readers directly to the URLs of open-access journal articles. None mentioned the Directory of Open Access Journals and other similar sites, even though about 37 percent of the referenced journals were indexed in the Directory. Why was there only a small fraction of open-access journals referenced? What are scholars' attitudes toward open-access journals? Do they consider these journals

to be as reputable as traditional print or electronic subscription-based journals? It is hoped that, by conducting a survey among scholars from different disciplines or employing a detailed analysis approach to the scholars' acceptance of open-access journals, future studies can provide a clearer picture of the role that open-access journals play in the academic communication system. The finding that the other two types of open repositories were heavily cited suggests that digital collections of scholarship are making a critically important contribution to the scientific community with regard to the use and dissemination of intellectual output. Additional study on open repository resources will reveal the factors that led to the high proportion of referrals to such resources. Noticeably, in the open-access repository movement, academic libraries have crucial influence on the building and delivery of repository services. The development of institutional repositories demands library technical services departments to develop a new set of management strategies to address collection, preservation, cataloging, and dissemination issues in the digital environment. The movement also calls for public services librarians to establish active communication with scholars, engage and work with them, and promote repositories to wider communities. Last, citations to numerous other kinds of Web-based materials imply that scholars do need to consult a broad range of conveniently accessible resources to meet their specific research needs. This encourages information professionals to develop strategies for the presentation and advocacy of reliable Web-based research publications.

Limitations of the Study

The present research has several limitations that need to be taken into account when generalizing the findings. First, the journal samples can include certain limitations. The study selected only ISI-indexed journals that have high-impact factors and that are subscribed to by the researcher's library. The results may have been differ-

ent for lower-impact journals that have fewer subscribers and, therefore, fewer chances to be formally cited. The study has focused on six subjects, representing a selection from the sciences, social sciences, and humanities disciplines. However, each discipline, even in the same general classification category, has its unique features. There could have been completely different findings revealed by examining other disciplines. Likewise, the study only covered Web references in the journal literature. Further investigation on Web citation behavior in monographs, book chapters, and conference presentation papers should be performed to complete a full analysis of the impact of Web resources on scholarly research. This paper, therefore, presents some information on the Web reference practices of scholars from one aspect and brings up possible avenues for future research.

Conclusions

The current study contributes to a small number of studies that examine the citing behavior of scholars concerning the impact of Web-based resources. The purpose is to discover the influence of the Web-based resources on research practices, identify the similarities and differences across multiple disciplines in incorporating Web references into the journal literature, and explore whether any changes emerged from 2001 to 2007. The results indicate that, in general, scholars cite relatively few Web resources. This phenomenon is particularly reflected in some of the natural science fields such as chemistry and biology. Many issues could lead to the low rate of Web references. Apart from the nature of different disciplines and the factor that scholars are likely using the electronic versions of journals but citing as if it is the print version, overall, the lack of quality control and the lack of persistence of Web references are perhaps the major hurdles that prevent scholars from taking advantage of the Web's greatest feature: providing immediate access to various scientific

materials. The study did not carry out an exhaustive investigation on all possible factors that may affect scholars when using the Web to share research findings. Future research may analyze the role that age or gender plays to provide more evidence on the patterns of Web-based information use. On the other hand, this research work echoes other studies that, even though there are not a great deal of Web information resources found in peer-reviewed journal articles, the number of such resources and the number of articles containing Web-based resources have noticeably increased between 2001 and 2007. Among the findings, the most interesting discovery is that open-access data repositories and open-source analysis programs or software are cited predominantly in the journal literature across the six chosen disciplines. This suggests the significance for information professionals to work on new strategies to manage Web-based resources and provide reliable, systematic repositories for supporting research and scholarship.

Notes

1. Stephen P. Harter and Hak Joon Kim, "Electronic Journals and Scholarly Communication: A Citation and Reference Study," *The Digital Revolution: ASIS Mid-Year 1996: Proceedings of the ASIS Mid-Year Meeting, San Diego, California, May 18–22, 1996* (Medford, N.J.: Information Today Inc., 1996): 299–315.

2. Yin Zhang, "Scholarly Use of Internet-based Electronic Resources," *Journal of the American Society for Information Science and Technology* 52, no. 8 (June 2001): 628–54.

3. Susan Davis Herring, "Use of Electronic Resources in Scholarly Electronic Journals: A Citation Analysis," *College and Research Libraries* 63, no. 4 (July 2002): 334–40.

4. Mary Rumsey, "Runaway Train: Problems of Permanence, Accessibility, and Stability in the Use of Web Sources in Law Review Citations," *Law Library Journal* 94, no.1 (Winter 2002): 27–39.

5. Paul Wouters and Repke de Vries, "Formally Citing the Web," *Journal of the American Society for Information Science and Technology* 55, no.14 (Aug. 2004): 1250–60.

6. Lyn Robinson, "Impact of Digital Information Resources in the Toxicology Literature," *Aslib Proceedings* 59, no. 4/5 (2007): 342–51.

7. Cecelia Brown, "The E-volution of Preprints in the Scholarly Communication of Physicists and Astronomers," *Journal of the American Society for Information Science and Technology* 52, no.3 (Feb. 2001):187–200.

8. ———, "The Role of Web-Based Information in the Scholarly Communication of Chemists: Citation and Content Analyses of American Chemical Society Journals," *Journal of the American Society for Information Science and Technology* 58, no.13 (Nov. 2007): 2055–65.

9. Henk F. Moed, *Citation Analysis in Research Evaluation* (Dordrecht, the Netherlands: Springer, 2005), 119–32.

10. Mu-hsuan Huang and Yu-wei Chang, "Characteristics of Research Output in Social Sciences and Humanities: From a Research Evaluation Perspective," *Journal of the American Society for Information Science and Technology* 59, no.11 (June 2008): 1819–28.

11. The following 30 journals were selected for analysis: Chemistry—*Analytic Chemistry, Inorganic Chemistry, The Journal of Organic Chemistry, Advanced Materials, Journal of the American Chemical Society*; Biology—*The FASEB Journal, Genes & Development, Evolution, Molecular Microbiology, Cell*; Economics—*Journal of Political Economy, Quarterly Journal of Economics, Journal of Financial Economics, Econometrica, The Review of Economic Studies*; Anthropology—*Journal of Human Evolution, Current Anthropology, American Journal of Physical Anthropology, American Journal of Human Biology, American Antiquity*; History—*American Historical Review, Journal of American History, Past & Present, Journal of African History, Journal of Modern History*; Linguistics—*Brain and Language, Journal of Memory and Language, Language, Linguistic Inquiry, Journal of Speech Language and Hearing Research.*

12. Internet Archive, "About the Internet Archive." Available online at www.archive.org/about/about.php. [Accessed 3 November 2009].

13. Helen Georgas and John Cullars, "A Citation Study of the Characteristics of the Linguistics Literature," *College & Research Libraries* 66, no. 6 (Nov. 2005): 496–515.

14. Diomidis Spinellis, "The Decay and Failures of Web References," *Communications of the ACM* 46, no. 1 (Jan. 2003): 71–77.

15. Wayback Machine. Available online at www.archive.org/web/web.php. [Accessed 3 November 2009].

acrl 2011

acrl 2011

peer revered

ACRL 2011

ACRL 2011 is the premier event for academic libraries, drawing librarians, support staff, and vendors from across the country and around the world. Register by **February 4, 2011**, and **save more than 20 percent**. Group rates, reduced airfares, hotel discounts, and scholarships are also available.

ACRL 2011 will feature award-winning filmmaker, artist, Internet pioneer, and activist **Tiffany Shlain**; activist, academic, and much-in-demand speaker, **Raj Patel**; and author, speaker, and co-host of the immensely popular TLC show *What Not to Wear*, **Clinton Kelly,** as well as **more than 300 peer-reviewed sessions**.


Jaron Lanier is a computer scientist, composer, visual artist, and author who writes on numerous topics, including high-technology business, the social impact of technology, the philosophy of consciousness and information, Internet politics, and the future of humanism. Don't miss his Saturday morning talk entitled, *The Bipolar Library: How Humanizing and Digitizing Must Both be Advanced*.

Just Added!

Jaron Lanier
Photo courtesy Jonathan Sprague
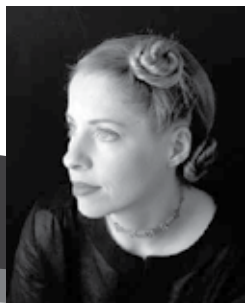
Tiffany Shlain

Raj Patel

Clinton Kelly

Connect online at **www.acrl.org/acrlconference**
www.facebook.com/acrl2011 | Twitter: #acrl2011