

Accessing Electronic Journals and Other E-publications: An Empirical Study

Stephen P. Harter and Hak Joon Kim

This article reports the results of a study of access problems and issues related to scholarly and peer-reviewed electronic journals (e-journals) currently found on the Internet. Data are reported on the accuracy of directory information, the accessibility of e-journal articles, the status of e-journal archives, methods used to retrieve current and back issues, and the variety of data formats used by e-journals. Data also are presented on the accessibility of the texts of electronic publications referenced by e-journal articles. The findings illustrate the practical problems that can arise when users attempt to retrieve the texts of electronic publications, and reveal relatively poor accessibility and usability of the e-journals studied. Implications for academic and research libraries are considered.



The first scholarly journals, *Journal des Sçavans* and the *Philosophical Transactions of the Royal Society*, began publication in 1665.¹ For more than three centuries, the journal has played a crucial role in the creation, transmission, and storage of knowledge, as the primary medium of communication among scholars. Despite its successes, however, in the past fifty years the print journal system has been subjected to many criticisms, perhaps the most serious of which are high costs, long delays in the publication process, and failures of peer review.² The problems of print journals and the concomitant development of computer and communications technology have led to the development of electronic alternatives to print

journals—various forms of the electronic journal, or e-journal.

This article reports empirical research conducted in the latter half of 1995 on access problems and issues related to networked e-journals, defined as e-journals that are delivered via, or accessed through, the Internet.³ It is part of a larger study designed to assess the impact of e-journals on scholarship and research by examining references of, and citations to, e-journal articles.⁴

The Research Problem

Before the introduction of the first graphical World Wide Web (WWW) browser (Mosaic) in 1993, most e-journals were based on the listserv (electronic mailing

Stephen P. Harter is a Professor in the School of Library and Information Science, Indiana University; e-mail: harter@indiana.edu. Hak Joon Kim is a Doctoral Student at the School of Library and Information Science, Indiana University; e-mail: hajkim@indiana.edu.

list) technology operating on the Internet and Bitnet networks and distributing e-journal articles as ASCII text files. The Gopher or FTP protocols were often used as complementary mechanisms providing additional access to listserv-based e-journals. Since 1993, the tremendous growth of the WWW has influenced both the number and distribution mechanisms of e-journals. Many e-journals formerly based on listservs and most newly established e-journals have migrated to the WWW, which allows journals to go far beyond simple ASCII text in the provision of information. Journals published on the WWW can include hypertext links connecting references, tables, and other parts of an article, as well as links to other Web files. In addition, they can use powerful new features such as graphics, sound, video, and even miniprograms (e.g., Java applets). Web browsers use several Internet protocols, including Gopher, FTP, and http (hypertext transfer protocol). In addition, speed of publication can be greatly enhanced over that of print journals. Thus, although thorny cost, quality, archival, and preservation issues have not yet been resolved, there are clearly potential benefits of e-journals over traditional print journals. Some commentators expect e-journals to supplant paper journals in the relatively near future.⁵

To this point, vastly more speculation has been written about the potential of e-journals than empirical research conducted on existing ones. Although thousands of e-journals, newsletters, zines, and other electronic serials now are available on the networks and interest in them has greatly heightened, little empirical research has been conducted.⁶

Hazel Woodward and Cliff McKnight argue that three levels of access to e-journals must be considered.⁷ The first is access to the bibliographic information on e-journals and publishers, including names and addresses. The second level

is access through indexing and abstracting services to information about the individual articles published by e-journals. The third level is access to the actual texts of e-journal articles.

Several e-journal directories are available in both print and electronic form,⁸ and some abstracting and indexing services have recently begun to index a few e-journals.⁹ However, to the authors' knowledge, no empirical research has been published regarding access problems and issues related to the texts of e-journal articles that are currently avail-

To the authors' knowledge, no empirical research has been published regarding access problems and issues related to the texts of e-journal articles that are currently available on the networks.

able on the networks. Reliable access and usability seem crucial if e-journals are to succeed as a new medium of formal scholarly communication.

A related issue is access to the texts of e-journal articles and other kinds of e-publications that have been cited in print or electronic journals. That is, when an author cites an e-publication, what can be said about the accessibility of that item, given the information provided in the citation? Again, to the authors' knowledge, no research has been conducted on this question.

The specific research questions addressed in this article include:

- To what extent can information provided in e-journal directories be used to retrieve the actual texts of articles? The issues addressed here are accuracy, currency, and completeness of the information provided, as well as stability of the locations of these files.
- If e-journals cannot be located from the information provided in a directory entry, can they be found through independent means (e.g., use of search engines)?

- What kinds of problems do users encounter when they try to access e-journals?

- To what extent do e-journal archives exist?

- How complete are existing e-journal archives?

- What methods are used to access current issues and archives of e-journals?

- What data formats are used for e-journals? These will have implications for ease of use, printability, and so on.

As part of the larger study, references in e-journal articles were checked as a measure of the impact of e-journals on scholarly communication. References to e-journal articles and other types of e-publications also were checked for accessibility and type of resource. Thus, two additional questions are raised:

- What are the frequencies and types of e-publications that are references in e-journal articles?

- To what extent are references to e-publications in e-journal articles accessible to users, based on information provided in the reference?

Methodology

The primary purpose of the larger study was to assess the impact of scholarly and peer-reviewed (or refereed) e-journals on scholarly communication by studying citations to, and references in, e-journal articles as of the latter part of 1995. To conduct the reference study, it was necessary to retrieve and study a sample of e-journal articles. This gave the authors an opportunity to experience firsthand the kinds of problems that occur when users try to accomplish the same task.

Because the larger study dealt with scholarly and peer-reviewed e-journals, the authors needed to exclude other electronic publications such as newsletters and zines. To define their population, the authors chose to use two print directories that include information on electronic journals: *Internet World's on Internet 94*, published by Mecklermedia, and *Dir-*

tory of Electronic Journals, Newsletters and Academic Discussion Lists, published by the Association of Research Libraries (ARL).¹⁰ All e-journals identified in either directory as peer-reviewed or refereed were included. Procedures followed to select the sample (drawn in June 1995) were:

- E-journals selected from the ARL directory were taken from the section entitled "Electronic journals, magazines, and zines" and identified there as "peer-reviewed." All e-journals meeting these criteria became members of the sample.

- E-journals selected from the Mecklermedia directory were taken from the section entitled "Electronic Journals and Newsletters" and identified there as "refereed." All e-journals meeting these criteria became members of the sample.¹¹

- The two lists were combined and duplicates removed. There were 131 scholarly and peer-reviewed or refereed e-journals in the final list. The authors attempted to access all of these to retrieve a sample of articles for study.

Most of the e-journals in the sample use more than one access mechanism. The authors used *all* access methods provided by the ARL and Mecklermedia directories to attempt to gain access to the e-journal collections. For example, one e-journal directory entry provided only the editor's e-mail address as an access method, whereas another provided the editor's e-mail address, uniform resource locators (URLs) for FTP, Gopher, and WWW, and a listserv address. The authors checked all the access points provided in each e-journal entry in the two directories for each of the 131 e-journals in the sample.

There were several occasions when the authors failed to gain access to the e-journals using any of the access methods provided. In such cases the authors used other methods. First, they used Internet search engines such as Infoseek, Webcrawler, Lycos, and Yahoo. In this way, some of the inaccessible e-journals were located. Second, the authors checked vari-

ous e-journal collection sites such as the one maintained by the Committee on Institutional Cooperation (CIC) in its Electronic Journals Collection.¹² Third, the authors attempted to browse (navigate) in directories of the host server, based on partial addresses provided in the directories. Although this technique is intuitive and time-consuming, it was generally useful because many of the inaccessible e-journals had incomplete URLs in their directory listings. Finally, when the authors failed to access an e-journal using all the techniques described above, they attempted to contact its editor. Sometimes this method also was successful.

Through their best efforts to access all the e-journals in the sample, the authors identified the extent to which e-journals were accessible, the types of problems that actually occur when users try to access e-journals, and some possible causes of these problems. In addition, the authors investigated whether archives for the e-journal existed and, if so, their state of completeness. Finally, the authors attempted to determine the life status of the e-journals in the sample. An e-journal was operationally defined as "alive" if its archives existed and contained at least one issue or article published in 1995 (as of September 1, 1995). Otherwise, the authors assumed it had ceased publication.

Findings

Access Methods and Data Formats

Table 1 reports the most frequently used combinations of methods for accessing e-journals. Most e-journals use two or more access/distribution methods. Twenty-four e-journals used the WWW as the sole access/distribution method, making it the most popular e-journal distribution mechanism. Six e-journals used Gopher alone. This is not a surprising result when the explosive growth of the WWW and the simultaneous drop in attention to Gopher, which itself underwent explosive growth only a few years earlier, are considered.

TABLE 1
Most Frequently Used Combinations of Methods for Accessing E-journals

Combination of Methods	No. of E-journals
World Wide Web (WWW) alone	24
Listserv, FTP, Gopher, and WWW	9
Listserv and WWW	8
Listserv, FTP, and Gopher	7
Listserv, FTP, and WWW	7
Listserv and Gopher	6
FTP, Gopher, WWW, and e-mail	6
Gopher alone	6

Table 1 also shows the heavy use of listserv software in combination with other distribution methods. Most e-journal publishers evidently recognize the limitations as well as the strengths of listserv-based systems and try to complement them with other methods. All listserv systems maintain a mailing list of subscribers. The table of contents and abstracts of an issue, or sometimes even the full texts of e-journal articles, are sent through e-mail whenever a new issue or article becomes available. If an e-journal is available only through the listserv, accessing e-journal issues, especially back issues, is somewhat difficult, particularly for non-subscribers.¹³ The procedures can be time-consuming and require a fair amount of knowledge on the part of users. On the other hand, to be notified of current issues or articles via e-mail is probably the most convenient of all approaches. Thus, it is not surprising to find listserv software used in combination with other distribution methods.

Accepting the premise that a major function of scholarly e-journals is "to make public the results of original research to the widest possible audience,"¹⁴ it follows that several methods should be used to provide access so that users can choose the methods most suitable for them. Nine e-journals use all four of the most common e-journal distribution methods: listserv, FTP, Gopher, and WWW.

Set your information



Scan any bound document
face up and damage-free
in seconds.



Introducing Minolta

Scan, copy and send information from any bound document interoffice, interlibrary or internationally. Bound up by bound documents? Digitize and distribute information worldwide with one of the EPIC 3000 system. Scan fragile books, documents and even 3-D objects face up and damage-free. Then archive or distribute the information however you please. EPIC 3000 for Windows

© 1996 MINOLTA CORPORATION Encourages full compliance with the U.S. statutory law that prohibits the copying

free!

Copiers
Cameras
Faxes
Digital Systems
Document Imaging
Camcorders
Binoculars
Color Sensors

Only from the
of Minolta



MINOLTA

EPIC 3000

lets you print or fax. E-mail information.
Store on any form of media. Or put it on the Internet.
The patron-friendly EPIC 3000 with touch screen
comes ready for paper or electronic output.
Don't get caught in a bound document bind.

For more information call **1-800-9-MINOLTA**

Simple touch screen monitor lets
patrons print or store to disk.



TABLE 2
Access Methods for Current and Archival
Issues of E-journals (N=125)*

Access Method	No. of E-journals Using Method	Percentage of Total
WWW	83	66.4%
Gopher	53	42.4
FTP	52	41.6
Listserv	48	38.4
E-mail	21	16.8
Guidon	7	5.6
Paper (for current issue only)	3	2.4
Telnet to freenet	1	0.8

*A given e-journal may provide several access methods.

Table 2 shows that about two-thirds of the e-journals in the sample used the WWW as at least one distribution method. Gopher, FTP, and listserv occupied the second, third, and fourth ranks, respectively. In addition, it should be noted that e-mail, which is relatively primitive compared to other e-journal distribution methods, is still used as a distribution method by one-sixth of the e-journals studied. Table 3 lists the very diverse data formats used by the e-journals in the sample. Among these, the simple ASCII text file was the most commonly used. ASCII is limited in terms of its information representation ability because it cannot handle graphics or special characters such as mathematical symbols. On the other hand, for certain scholarly forms (e.g., essays) and for the bulk of research and scholarship in many fields (e.g., English and sociology), ASCII is perfectly adequate and perhaps preferable. Plain ASCII texts also have an obvious merit in that they can be easily read and manipulated by all kinds of computer hardware and software,

and can easily be transmitted using the straightforward e-mail and listserv technologies. HTML (Hypertext Markup Language) and PostScript are the second and third most commonly used data formats. HTML is the standard language used on the WWW for creating and displaying hypermedia documents. PostScript is Adobe's page description language used for delivering and printing complex documents that include layout, font descriptions, and graphics over the Internet.

Table 3 also shows that various data formats necessary for multimedia presentations, such as DVI, GIF, MPEG, and QuickTime, have begun to be

TABLE 3
Data Formats Used by E-journals*

Data Format	No. of E-journals
Adobe Acrobat PDF (Portable Document Format)	6
ASCII (American Standard Code for Information Interchange)	80
DVI (Digital Video Interactive)	6
GIF (Graphics Interchange Format)	6
Guidon	7
HTML (Hypertext Markup Language)	62
Hypercard	2
hyperTeXdvi	1
JPEG (Joint Photographic Experts Group)	2
LaTeX	2
Microsoft Word	2
MPEG (Moving Pictures Experts Group)	1
PostScript	24
QuickTime	1
RTF (Rich Text Format)	4
TeX	8
TIF (Tagged Image File)	1
Word Perfect	3

*A given e-journal may provide data in several data formats.

TABLE 4
Success Rates in Connecting
to E-journal Archives*

Connection Result	No. of E-journals	Percentage of Total
Connected first time for all methods given in the directories	58	45.0%
Failure to connect for one or more methods	71	55.0

*Two fee-based e-journals could not be accessed because the authors had no subscription.

used by e-journals. The kinds and numbers of data formats used for multimedia presentation indirectly reflect the impact of the WWW on e-journals. With the exception of ASCII text, all the data formats listed in table 3 require special software, and in some cases hardware, to view or use the document or document part. This fact can have a significant impact on user accessibility (and on a library's ability to provide access for users to these files).

Accessibility of E-journal Articles

Table 4 shows the success rates the authors achieved in connecting to the e-journal archives using all access methods listed in the ARL and Mecklermedia directories. Less than half the e-journals in the sample could be accessed on the first try using all the methods and addresses provided. This is a disappointing result, because one of the most important advantages claimed for e-journals compared to print journals is speed of access. Such low success rates in accessing e-journals must be considered a serious

problem. If e-journal articles cannot be accessed by the target audience when they are needed, their value is diminished. Unstable or limited access to e-journals cannot help but affect their acceptability among scholars as a legitimate medium of formal scholarly communication.

Table 5 reports the results of the authors' failures to connect to e-journal archives. More than 40 percent of the e-journals in the sample had one or more identifiable problems with the data provided in one or both directories.

Nearly 10 percent of the e-journals could not be accessed at all through the URLs listed, although the authors tried to connect to each on at least three different days and times. Eleven (8.6%) of the e-journals had moved their locations, but only six of these provided a "move notice" on their old sites that included a new address (URL). Five e-journals moved but provided no information. In such cases, the authors were able to identify the new location through their own efforts, for example, by using Internet search engines or via intuitive navigation of host serv-

TABLE 5
Summary of Failures in Attempts
to Connect to Archives*

Type of Problem	No. of E-journals	Percentage of Total
Identifiable problems in directory data	55	42.6%
No response to listserv or e-mail request	9	7.0
Failure to connect to Gopher URL on three or more different days	6	4.7
Location has moved, new address given	6	4.7
Location has moved, no new address given	5	3.9
Failure to connect to Web URL on three or more different days	3	2.3
Failure to connect to FTP URL on three or more different days	3	2.3
Listserv responded but did not accept subscription to e-journal	1	0.8

*More than one result is possible for a given e-journal. N=71.

TABLE 6
Problems with Data Provided in
the Print Directories*

Type of Problem	No. of Times Problem Occurred
Part of URL was missing	20
Missing URL	14
Typographical error	9
Incorrect listserv command given	9
URL was entirely incorrect	9
Part of URL was incorrect	3
Incorrect e-mail address	2

* More than one problem may be associated with a given e-journal.

ers. In addition, nine (7%) of the e-journals did not respond in any way to the authors' listserv or e-mail requests. For these e-journals the authors can conjecture that either they have ceased publication, no longer distribute issues using e-mail, or have a new editor (with the former editor no longer answering his or her e-mail).

Nearly half the problems the authors experienced in connecting to the archives of the e-journals in the sample could be traced to inaccurate data in the print directories (see table 6). For twenty e-journals, parts of the URLs were missing. The authors were often able to identify the correct URLs by navigating the sites using the partial information provided in the directory entry.

In fourteen cases, the authors discovered a Gopher, FTP, or Web location for an e-journal through their own efforts, but this information was not provided in the relevant directory entry. These situations almost certainly resulted from the rapidity of change—from an e-journal adding a Gopher or Web site after the directory information had been gathered and published.

The authors found nine typographical errors in the e-journal listings. The types of errors varied

from relatively easy to find to very difficult. For example, "SUBSCRIBED" instead of "SUBSCRIBE" was given as a listserv command, a relatively easy error to diagnose (but not, perhaps, for a naive end user). In another case, "utas" instead of "utias" was given as a part of a URL; this was very difficult to detect. Even though these look like very minor errors, access to the e-journals failed because of them. URLs must be correct in every aspect, including case, or the connection will fail. "Minor" typographical errors can have devastating consequences.

In nine directory entries, an incorrect listserv command was provided. For example, one of the directories gave the subscription command syntax "subscribe <list>" instead of "subscribe <list> <your name>" for subscribing to a list controlled by the Listproc listserv. The command <your name> is necessary in Listproc, but optional in Listserv.

In nine additional cases, the URL provided was entirely incorrect, in many cases probably due to the archive having been moved to a completely different location. Other less frequently observed problems in the directory listings are given in table 6.

Another important issue that should be considered in determining the accessibility of e-journals is the completeness

TABLE 7
Completeness of Archives

State of Archives	No. of E-journals	Percentage of Total
Archives were complete for all methods	87	66.4%
Archives were incomplete for one or more methods	28	21.4
Impossible to evaluate completeness of archives	16	12.2
Total	131	100%

of the archives. Many of the e-journals provide several different access methods, and also maintain several different forms of archives based on these methods. Table 7 reports the authors' findings on the completeness of the archives maintained by each e-journal in the sample. For two-thirds of the e-journals, all archives were complete. More than one-fifth of the e-journals had at least one incomplete archive, and the authors could not determine the completeness of the archives of the remaining sixteen. To publish their articles as quickly as possible, some e-journals have no concept of issue, though they usually retain the concepts of volume and/or publication year. For these e-journals, articles are individually distributed (published) as they are accepted. Such a mechanism, which is unique to e-journals, speeds up the publication process because there is no need to wait until enough articles are ready before publishing an issue. However, this made it impossible to identify whether articles were missing from the e-journal archives, because typically no index to published articles was provided.

Table 8 reports the life status of the e-journals in the sample. About a quarter of the e-journals had ceased publication, according to the authors' operational definition of *death*. Twenty-eight e-journals had archives but had published no 1995 issue (or article). Four e-journals could not be accessed at all. In addition, the authors could not determine the life status of four e-journals. Two of these were fee-based e-journals for which the authors had no subscription. The remaining two had no concept of volume, issue, or publication year. Although there were articles stored in the archives, it was impossible to tell when they had been placed there.

E-journal Pricing and Access
Whether an e-journal charges a

subscription fee or is free can affect the extent to which libraries may be asked for assistance in providing access to it. The authors identified eight pricing models for the e-journals in the sample (see table 9). In all, only thirteen (10.4%) of the 125 e-journals in the sample for which the fee question could be determined charged a subscription fee.

The authors tested the relationship between several access variables and whether a subscription fee was charged, using the chi-square test. A significant relationship was found between the success the authors had in connecting to the e-journal archives (see table 4) and whether a subscription fee was charged. Free e-journals were significantly more likely to have access problems ($\chi^2 = 11.172$, $p = .002$).

References to E-publications

One of the main purposes of the larger study was to analyze the references from a sample of articles in scholarly and peer-reviewed e-journals in order to measure the extent to which authors of e-journal articles are currently citing e-journals and other online sources. Such authors can be assumed to be more knowledgeable about, and sympathetic to, the electronic media than typical authors. Thus, the re-

TABLE 8
Life Status of E-journals in the Sample

Life Status	No. of E-journals	Percentage of Total
Archives and one issue in 1995 exist [living]	95	72.5%
Archives exist but no issue published in 1995 [probably ceased publication]	28	21.4
No connection could be made [probably ceased publication]	4	3.1
Could not determine life status	4	3.1
Total	131	100%

TABLE 9
Pricing Arrangements for E-journals

<i>If There Is No Print Counterpart (N=101)</i>		
Pricing Arrangement	No. of E-journals	Percentage in Category
All electronic issues are free	91	90.0%
Fee for both archival and current issues	8	7.9
Fee for current issue; archival issues are free	1	1.0
Subscription to electronic version is included with membership in a society	1	1.0
<i>If There Is a Print Counterpart (N=24)</i>		
Pricing Arrangement	No. of E-journals	Percentage in Category
Electronic version is free; fee for print version	16	66.7%
Fee for both electronic and print versions	4	16.7
Electronic version is free; cannot tell about print version	2	8.3
Electronic version is free; subscription to print version is included with membership in a society	2	8.3
Total (fee charged)	13	10.4%
Total (free)	112	89.6%

sults of the reference study should provide an upper bound on the current influence of e-journals on scholarly communication.

Among the original sample of 131 e-journals, the authors determined through examination of sample issues that seventy-four of these were scholarly and peer-reviewed, were accessible to them, and published articles with references. The authors retrieved and printed the last four such articles published by each e-journal.¹⁵ Six of the seventy-four e-journals had published fewer than four articles as of September 1995, when the data collection took place. In these cases the authors studied all the available articles.

From the most recent 279 articles published in the seventy-four scholarly and peer-reviewed e-journals, there were a total of 4,317 references. The authors classified each reference as to its format

(book, serial, book chapter, online source, etc.). They further classified the online sources as to subtype (Web page, e-mail, e-journal article, etc.). Finally, the authors tried to retrieve the texts of the referenced online sources. A full discussion of the methodology and findings from the reference study is given in Harter and Kim.¹⁶ Of particular interest here is the result of the authors' attempts to access the online publications that were cited in the sample of e-journal articles.

A total of eighty-three online sources (1.9% of the 4,317 references) were cited. Table 10 shows the types and frequencies of references to online sources, arranged by the content of the source (Web page, listserv posting, etc.). Web pages, electronic personal papers, e-mail messages, and e-journal articles were the most commonly referenced electronic sources. The authors found many of the references to

TABLE 10
Types, Number, and Accessibility of Cited Online Sources

Type of Online Source	No. of References	No. and Percentage of Accessible* References
Web Page	12	10 (83.3%)
Electronic Personal Paper	10	7 (70.0)
E-mail Message	9	2 (22.2)
E-journal Article	9	6 (66.7)
Newsgroup Posting	7	0 (0.0)
Listserv Posting	5	2 (40.0)
Electronic Directory	5	5 (100.0)
SGML Encoded Document	5	5 (100.0)
Electronic Preprint	3	1 (33.3)
Computer Software	2	2 (100.0)
Electronic Newspaper	2	2 (100.0)
Online Catalog	1	1 (100.0)
Local File	1	0 (0.0)
Type could not be determined	12	0 (0.0)
Total	83	43 (51.8)

*A reference is accessible if it leads directly to the text cited given only the information found in the citation.

online sources to be inconsistent, incomplete, and/or inaccessible. Many did not lead to the text of the cited online source.

The authors could not determine the types of twelve online references (14.5%) because bibliographic information was incomplete and because the online sources were not accessible, making further investigation impossible. For example, some of the online references provide *only* a URL and no other bibliographic information. In these cases, if the authors could not access the cited reference through the given URL, it was usually not possible to determine the type of online source being cited.

The authors studied each of the eighty-three references and tried to access the text of the cited material. One of the great theoretical advantages of references to online sources is that the reader can retrieve the full texts of the cited sources very quickly (ideally, by clicking on the reference itself). The third column of table 10 shows the results of our attempts to retrieve the texts cited by the eighty-three

references. Only six of the nine e-journal articles cited were accessible. E-mail messages and newsgroup postings were especially inaccessible, and none of the twelve sources whose types could not be determined could be retrieved. Overall, only about half of the texts of the cited online sources could be accessed.

These findings are analyzed in another way in table 11. Online information sources can be obtained using several different access protocols. Forty-seven of the cited references included a URL. The authors used Netscape, a Web browser, to attempt to access these URLs. However, only two-thirds of the URLs led to the text of the source. For the remaining third, the authors obtained the browser message "The server does not have a DNS entry" or the server error message "404 URL Not Found." Several types of the online references listed in table 10 (e.g., Web pages, electronic personal papers, SGML-encoded documents, software programs, and local file references) were usually cited in this way, often *only* with a URL.

TABLE 11
Access Protocols and Accessibility for
the 83 Online References

Access Protocol	No. of Online References	No. and % of Accessible References
URL	47	31 (66.0%)
E-mail information	13	4 (30.8)
Listserv information	9	6 (66.7)
Usenet newsgroup	7	0 (0.0)
Electronic newspaper	2	2 (100.0)
Incomplete access information provided	5	0 (0.0)
Total	83	43 (51.8%)

For online sources in which e-mail information was provided, the authors tried to contact the source using e-mail. Only four of these references led to the online source cited. The authors received the mail server message "User Unknown" in seven cases. In two other instances, there was no response to the authors' message.

Nine of the online references were to listserv postings. The authors were able to access six of these by using the listserv access information provided in the reference (e.g., the listserv name and posting date) and by using listserv commands. These required a fair amount of knowledge and effort. The authors' failures were due to the lack of the existence of listserv archives or archives for the time period in question. This also was the explanation for the authors' inability to find the seven cited Usenet newsgroup postings. Two electronic newspaper articles were cited. The authors were fortunate to have access to Lexis/Nexis, and found the cited articles that way.

Finally, table 11 indicates that five online references fell in the category "incomplete access information provided." In these cases, there was insufficient information for the authors even to attempt to gain access to the source.

Table 11 shows that, except for electronic newspapers available through a

commercial service (to which not all Internet users will have access), none of the methods led to perfect retrieval. The authors consider the overall rate of 51.8 percent to be extremely low, given that the online sources were apparently considered important enough for an author to cite; and yet in just one or two years, nearly half of the referenced texts were not accessible online. The percentage of successfully accessed material would be

even lower if the authors required *transparent accessibility*—that is, if texts found through search engines or through knowledge of arcane listserv commands were considered to be in a different category than those available simply by clicking on a URL. Table 12 summarizes the accessibility findings when the data are viewed in this way.

Table 12 shows that only 37.3 percent of the references in the sample provided transparent (clickable) access to the text of a cited reference. In another 14.5 percent of the references, the text could be found, but only with significant effort, as described above. Nearly half of the references did not lead to the cited text even when significant effort was expended.

Direct accessibility to cited e-journal articles and other cited online sources would obviously be extremely convenient to readers and would offer a great advantage to e-journals over print. Equally, if readers cannot access original sources of referenced material, references to electronic sources are much *less* useful than references to print sources. The authors should emphasize that the references studied appeared in *recently published* e-journal articles—most during the year in which the data gathering took place. The authors might reasonably expect the great majority of such references to be largely correct and complete—at least enough to

lead the reader to the text of the referenced source. This was not what the authors found; only half of the references were accessible, and only somewhat more than a third provided transparent access. When more time has passed, these proportions are likely to be much lower. The authors conclude that the accessibility of cited online sources is potentially a very serious problem for e-journals and for the citation of electronic publications more generally.

Conclusions

The findings of this study illustrate the practical problems that occur when users attempt to access the texts of articles published in scholarly and peer-reviewed e-journals that are currently available on the Internet. The problems the authors identified raise several important issues that libraries should consider.

If academic and research libraries are to provide users with access to the texts of e-journals, appropriate hardware and software is needed. Many of the e-journals in the sample employ two or more access and distribution methods, and some include data formats such as PostScript, DVI, and TeX, as well as audio and video capability, for rich information representation. Many users will not be able to fully utilize such e-journals on their own computers for lack of their own appropriate hardware or software. Just as academic libraries subscribe to print journals for those users who are not able to individually subscribe to many journals in their fields, they also will need to provide users with appropriate equipment and software necessary not only for accessing e-journals, but also for printing articles and manipulating related files. In addition, user training will be necessary for many users if they are to effectively use the facilities provided.

One of the troubling findings of this study was with the information found in the two (print) e-journal directories the authors used. More than half of the e-journals in the sample could not be accessed by one or more methods given in the directories, despite the recent publication dates of the directories. The most frequent reason for access failure was inaccurate directory information. The authors assume that the main cause of inaccuracies is that the directories, which are compiled and published annually, cannot keep up with rapid changes in the relevant information. As long as growth and change in e-journals continue to be rampant, the print medium may not be well suited for an e-journal directory. In fact, the ARL directory is available in both print and electronic form. However, the electronic version, which is available on a Gopher site,¹⁷ is an abridged version of the annual print e-journal directory and is not continually updated. In this sense, it does not take full advantage of its electronic character. To keep up with rapid changes in e-journals, up-to-date information should be readily available—and that implies that the directories themselves should be on the Internet and in a state of continuous revision.

Another important issue is the need to create and maintain e-journal archives. A full 55 percent of the e-jour-

TABLE 12
Summary of Accessibility Results for
References to E-publications

Result of Accessibility Tests	No. of References	% of References
Transparent (clickable) access to text of cited reference	31	37.3%
Text of cited reference findable with significant effort, given the information provided in the citation	12	14.5
Text of cited reference not findable given the information provided in the citation	40	48.2

nals in the sample could not be accessed on the first try by one or more methods listed in the directories. More than half of the access failures were due to the problems of the e-journals themselves, such as servers being down or server location changes. Moreover, about one-fifth of the e-journals had incomplete archives and almost a quarter had probably ceased publication. The authors' data raise serious questions concerning the permanence and reliability of e-journal archives.¹⁸

Commercial publishers also will become much more involved. This article has dealt with scholarly peer-reviewed e-journals that were in existence as of June 1995 (as given in the ARL and Mecklermedia directories used to draw the sample). These did not include *any* of the major commercial publishers, such as Elsevier Science, Springer Verlag, John Wiley & Sons, Blackwell Science Limited, Academic Press, and Taylor and Francis, who are just now making their presence known.¹⁹ The authors are sure that some of the access problems noted here will be

addressed by these large fee-based publishers, who will probably try to maintain complete and easily accessible archives. The situation is not so clear for the smaller publishers. Indeed, the authors found significantly more access problems with free e-journals than with fee-based ones. This implies that smaller e-journal publishers need to give more attention to quality control issues, especially those involving access. It also suggests that the idea of free e-journals may be an illusion that cannot be sustained.²⁰

Academic libraries have traditionally performed archival functions for serial publications that publish academically relevant articles, regardless of format. E-journals should not be an exception. Precisely how and in what format e-journals should be archived by academic and research libraries—paper, microfilm/fiche, floppy disk, CD-ROM, or on local servers—remains to be determined by the individual libraries or cooperative entities that determine to maintain archives. But without such archives, the unreliable access to e-journals that the authors have reported makes their evolution into a new medium of scholarly communication somewhat problematic.

The traditional scholarly journal system was not built in a day. The transformation of print scholarly journals to their present form and function has been an ongoing process for more than two centuries. The journal has had to continuously evolve in order to best serve its role as the primary medium of scholarly communication. On the other hand, the scholarly e-journals that began to be available on the networks in the late 1980s have had a relatively short history, and academic libraries have grappled with them for an even shorter period. E-journals are still in a very early stage of development, and the many issues related to them are in constant change. Nonetheless, e-journals have unique advantages over print journals and offer the potential to revolutionize the traditional scholarly communica-

Index to advertisers

ACRL	480, 492
Archival Products	427
Assoc. of Christian Libs.	483, 490
Biosis	414
Bowker	401
EBSCO	cover 2
Elsevier Science	458
Greenwood Publishing	491
ISI	478
Libraries Unlimited	404
Library Technologies	457
McGraw-Hill	439
Minolta	444-45
OCLC Online Computer	438, cover 3
PAIS	397
Readmore	476
Systems Research	398
Todd Enterprises	cover 4
H. W. Wilson	428

tion system that heretofore has been dominated by paper journals. The change in the scholarly communication system brought by e-journals has already begun and will probably be profound.

Electronic journals are, of course, only one type of e-publication, though certainly one of the most important ones from the perspective of a research library. The study of references to e-publications reported here shows that access to the texts of the cited material is no more reliable than access to e-journal articles. Potential solutions to this problem are not apparent because there are vastly more citable e-publications on the WWW that are *not* e-journal articles

than those that are. Building a framework to approach this knotty problem is extraordinarily challenging. Yet traditionally academic libraries have played a crucial role in the scholarly communication process, and cannot ignore the access problems and issues the authors have discussed. It is time to actively attack them.

The authors wish to thank Margaret Harter, Debora Shaw, and Taemin Park for their thoughtful comments and questions on an earlier version of this manuscript, and are grateful to OCLC, Inc., for partial funding of the larger study of e-journals of which this report is a part.

Notes

1. Charles B. Osburn, "The Place of the Journal in the Scholarly Communications System," *Library Resources & Technical Services* 28, no. 4 (Oct./Dec. 1984): 315-24.

2. For discussions of the traditional journal system, see Bernard Houghton, *Scientific Periodicals: Their Historical Development, Characteristics and Control* (Hamden, Conn.: Linnet Books, 1975); Fritz Machlup and Kenneth Leeson, *Information through the Printed Word: The Dissemination of Scholarly, Scientific, and Intellectual Knowledge, vol. 2, Journals* (New York: Praeger, 1978); and Donald W. King, Dennis D. McDonald, and Nancy K. Roderer, *Scientific Journals in the United States* (Stroudsburg, Penn.: Hutchinson Ross Publishing, 1981).

3. Woodward divides electronic journals into three types: online (available from a commercial vendor such as DIALOG or Lexis/Nexis), CD-ROM, and networked (available directly on the Internet or other academic network). See Hazel Woodward, "The Impact of Electronic Information on Serials Collection Management," *IFLA Journal* 20, no. 1 (1994): 35-45. However, because DIALOG and other vendors are now searchable on the Internet, these are no longer mutually exclusive categories. Operationally, networked e-journals are defined in this article by the inclusion criteria used by the two directories the authors used to draw their samples: Mecklermedia's *Internet World's on Internet 94* and ARL's *Directory of Electronic Journals, Newsletters and Academic Discussion Lists* (see note 8 for full citations). Neither directory includes e-journals available through a commercial vendor such as DIALOG.

4. Selected findings from the full study are found in Stephen P. Harter and Hak Joon Kim, "Electronic Journals and Scholarly Communication: A Citation and Reference Study" (paper presented at midyear meeting of the American Society for Information Science, San Diego, Calif., May 18-22, 1996). The text of the paper is available at URL: <<http://www-slis.lib.indiana.edu/PrePrints/harter-asis96midyear.html>>. The portion of the study presenting the results of several citation analyses of selected e-journals is reported in Stephen P. Harter, "The Impact of Electronic Journals on Scholarly Communication: A Citation Analysis," *Public-Access Computer Systems Review* 7 (1996) [in press]. PACS Review is available at URL: <<http://info.lib.uh.edu/pacsrev.html>>.

5. Andrew M. Odlyzko, "Tragic Loss or Good Riddance? The Impending Demise of Traditional Scholarly Journals," *International Journal of Human-Computer Studies* 42 (1995): 71-122; Stevan Harnad, "The Postguttenberg Galaxy: How to Get There from Here," *Information Society* 11 (Oct./Dec. 1995): 285-91.

6. For exceptions, see Blaise Cronin and Kara Overfelt, "E-journals and Tenure," *Journal of the American Society for Information Science* 46, no. 9 (Oct. 1995): 700-03; Hans Roes, "Electronic Journals: A Survey of the Literature and the Net" (Dec. 1994), available at URL: <<http://cwis.kub.nl/~dbi/cwis/pub/ejournal.htm>>; Steve Hitchcock, Leslie Carr, and Wendy Hall, "A Survey of STM Online Journals 1990-95: The Calm before the Storm" (Jan. 1996), available at

URL: <<http://journals.ecs.soton.ac.uk/survey/survey.html>>; and F. W. Lancaster, "Attitudes in Academia toward Feasibility and Desirability of Networked Scholarly Publishing," *Library Trends* 43 (spring 1995) 741-52.

7. Hazel Woodward and Cliff McKnight, "Electronic Journals: Issues of Access and Bibliographical Control," *Serials Review* 21, no. 2 (summer 1995): 71-78. A useful analysis of access options is provided by Robert D. Cameron, "To Link or to Copy? Four Principles for Materials Acquisition in Internet Electronic Libraries," available at URL: <<http://fas.sfu.ca/0/projects/ElectronicLibrary/project/papers/e-lib-links.html>>.

8. Tony Abbott, ed., *Internet World's on Internet 94: An International Guide to Electronic Journals, Newsletters, Texts, Discussion Lists, and Other Resources on the Internet* (Westport, Conn.: Mecklermedia, 1994); Ann Okerson, ed., *Directory of Electronic Journals, Newsletters and Academic Discussion Lists*, 5th ed. (Washington, D.C.: ARL, 1995) (a truncated version of this is available at URL: <gopher://arl.cni.org:70/00/scomm/edir/edir95/jz>); "NewJour: Electronic Journals and Newsletters" (the Internet list for new journals and newsletters available on the Internet), available at URL: <<http://gort.ucsd.edu/newjour/>>.

9. For example, the Institute for Scientific Information currently indexes four e-journals in its databases: *Chicago Journal of Theoretical Computer Science*, *Journal of Artificial Intelligence Research*, *Gene-Combis*, and *Online Journal of Knowledge Synthesis for Nursing*. Two additional titles are targeted for inclusion in the near future. Source: E-mail message from Eileen Curtin, Editorial Development, Institute for Scientific Information, to Stephen P. Harter, Feb. 12, 1996.

10. As of June 1995, when the study began, the 1995 edition of the Mecklermedia directory had not been published.

11. By inspecting the e-journals themselves, the authors later reduced this list to exclude many e-journals, zines, and newsletters that in their judgment did not report the results of research or scholarship, even though they identified themselves in one or both directories as peer-reviewed or refereed. The reference study was conducted on this smaller sample.

12. The CIC Electronic Journals Collection is available at URL: <<http://ejournals.cic.net/>>. Other collections are "NewJour: Electronic Journals & Newsletters," available at URL: <http://gort.ucsd.edu/newjour/>; and "Scholarly Journals Distributed via the World Wide Web," available at URL: <<http://info.lib.uh.edu/wj/webjour.html>>.

13. There are several listserv software packages, including Majordomo, Listproc, Mailbase, and Listserv. Both the available features and commands differ significantly among these packages. Some listserv software does not allow the archiving, searching, or retrieving of files. The commonly used and venerable Listserv software by Eric Thomas is fully featured and powerful, albeit somewhat difficult to use. First, a nonsubscriber must send a subscription command to the Listserv system and then wait (for varying lengths of time) until getting an acceptance message from the system. Some listserv systems require a confirmation response of the subscription request. After the acceptance message has been received, the newly subscribed user must send an "Index" command to identify the name of the available files. After likely file names are identified, he or she must send a "Get" command to obtain the file. In the present article, lowercase "listserv" refers to all the listserv software packages and uppercase "Listserv" refers to Eric Thomas's product.

14. Jill Lambert, *Scientific and Technical Journals* (London: Clive Bingley, 1985).

15. The authors are grateful to OCLC, Inc., and to the named e-journal publishers for providing temporary access to the following fee-based e-journals for purposes of conducting this study: *Applied Physics Letters Online*, *Current Opinions in Medicine*, *Current Opinions in Biology*, *Immunology Today Online*, *Online Journal of Current Clinical Trials*, and *Online Journal of Knowledge Synthesis for Nursing*.

16. Harter and Kim, "Electronic Journals and Scholarly Communication."

17. See note 8.

18. Citation in a networked information environment and establishing trusted channels of distribution are two of the "integrity issues" discussed by Clifford A. Lynch in "Integrity Issues in Electronic Publishing," in *Scholarly Publishing: The Electronic Frontier*, Robin P. Peek and Gregory B. Newby, eds. (Cambridge, Mass.: MIT Pr., 1996), 133-45.

19. See Gary Taubes, "Science Journals Go Wired," *Science* 271 (Feb. 9, 1996), 764-67. Also available at URL: <<http://science-mag.aaas.org/science/scripts/display/full/271/5250/764.html>>.

20. For a thorough discussion of many economic issues related to e-journals, see Ellen Finnie Duranceau, ed. (with contributions by Stevan Harnad, Frank Quinn, Gail McMillan, Fytton Rowland, Robert H. Marks, and Janet H. Fisher), "The Economics of Electronic Publishing," *Serials Review* 21 (spring 1995): 77-90.

*A New Service on the
Information Superhighway*

Authority Express

If you have been searching for an easy way to authority control your library's current cataloging, try LTI's *Authority Express* service.

With *Authority Express*, a library uses the Internet to transmit a file of newly cataloged bibliographic records to LTI (via FTP). Overnight, LTI processes the records through its state-of-the-art authority control system. Then, at the library's convenience, it logs into LTI's FTP server to retrieve fully authorized catalog records, along with linked LC name and subject authority records.

Authority Express

- Keeps authority control current at an affordable price
- Integrates easily into existing workflows
- Lowers cost by reducing staff time spent on catalog maintenance
- Provides next-day turn around for up to 5,000 catalog records
- Accepts records for processing even if LTI did not perform the original authority control

"Authority Control for the 21st Century"



LIBRARY TECHNOLOGIES, INC.

2300 Computer Avenue, Suite D-19 Willow Grove, PA 19090

(215) 830-9320 Fax: (215) 830-9422

(800) 795-9504 email: LTI@LibraryTech.Com

Announcing a New Major Reference Work

Editors-in-Chief:
Alan R. Katritzky, FRS, USA
Charles W. Rees, CBE, FRS, UK
Eric F.V. Scriven, USA

Also Available
on CD-ROM!

Comprehensive Heterocyclic Chemistry II

Due December 1996: This new set has two major aims: to report on the enormous advances in the subject in the last decade and to extend the coverage into areas where CHEC left significant gaps. The latter includes particularly saturated and partially saturated

compounds, rings containing less common heteroatoms, and many classes of bicyclic and polycyclic systems. Material included in CHEC is, for the most part, not repeated in CHEC II and the two works are designed to be used together.

**Free Prospectus Available—
Send for Your Copy Now!**

**For Your Free Prospectus
or More Information:**

- Volume 1:** Three- and four-membered rings, with all fused systems containing three and four membered rings
- Volume 2:** Five membered rings with one heteroatom and carbocyclic fused derivatives
- Volume 3:** Five-membered rings with two heteroatoms and carbocyclic derivatives
- Volume 4:** Five-membered rings with more than two heteroatoms and fused carbocyclic derivatives
- Volume 5:** Six-membered rings with one heteroatom and fused carbocyclic derivatives
- Volume 6:** Six membered rings with two or more heteroatoms and fused carbocyclic derivatives
- Volume 7:** Fused five- and six-membered rings without ring junction heteroatoms
- Volume 8:** Fused five- and six- membered rings with ring junction heteroatoms
- Volume 9:** Seven membered and larger rings and fused derivatives
- Volume 10:** Author and Ring Indexes
- Volume 11:** Subject Indexes

UK & ALL OTHER COUNTRIES:

Elsevier Science Ltd
The Boulevard,
Langford Lane,
Kidlington,
Oxford OX5 1GB
UK
tel. 44 (0) 1865-843685
fax 44 (0) 1865-843946

NORTH AMERICA:

Elsevier Science
Regional Sales Office
655 Ave of the
Americas
New York, NY 10010
USA
tel. 212-633-3730
fax 212-633-3680



PERGAMON

An imprint of Elsevier Science