

Availability and Half-life of Web References Cited in Information Research Journal: A Citation Study

A. Isfandyari Moghaddam, Ph.D.

Islamic Azad University
Hamedan Branch, I. R. of Iran
Corresponding Author:
ali.isfandyari@gmail.com

M. K. Saberi, M.A.

Islamic Azad University
Sciences and Research Branch, Tehran, I. R. of Iran
email: mohamadsaberi@gmail.com

S. Mohammad Esmaeel, Ph.D.

Islamic Azad University
Sciences and Research Branch, Tehran, I. R. of Iran
email: m.esmaeili@sr.iau.ir

Abstract

This article examines availability and half-life of URLs cited in articles published by *Information Research* journal. To do this, at first, we extracted all issues of *Information Research* from April 1995 to March 2008 and calculated number of all citations whether printed citations or Web ones. Afterwards, we checked availability of individual cited URLs. When we could not access directly URL inserted in any article by author(s), we tried to visit referred website. If this attempt seemed to be inadequate, search engine "Google" was employed to access the missing reference(s). Research findings indicated that 66% of articles have Web citations and rate of articles containing URL has increased from 17% in 1995 to 89% in 2008. Domains .net and .org have more stability and persistence compared to domains .edu, .gov, .uk and .com. Also, of 1761 cited URLs, 73% were accessible, and 27% were inaccessible. It is notable that using Google and searching missing URLs, accessible URLs increased from 73% to 86%. Finally, it was recommended that the best solution to prevent decay or disappearance of Web citations and diminish URLs decay is to make use of WebCite[®]-enhanced reference.

Keywords: Internet Research, Citation, Citation Analysis, Web Citation, Internet Resources Half-time, URL.

Introduction

Since the quasi-miraculous emergence of the Web in 1990s, there has been a continuous increase in the volume of scholarly resources in electronic form, such as e-books, e-journals, e-databases, e-theses and dissertations, e-prints of research papers, and the like. These resources have provided a scope for researchers and authors in various

subject fields and stimulated their research productivity. Library and information science (LIS) is no exception to this" (Maharana, Nayak & Sahu, 2006). Consequently, as a result of the invention of WWW, citation behavior of researchers has been influenced and in fact, it can be said that they have been interested in Web citations. As Zhao and Logan, 2002, have indicated, the reason behind such an increase in the number of Web citations in scholarly papers in LIS is that the Web has become the first choice for finding information on current research, for breaking scientific discoveries and for keeping up with colleagues at other institutions. On the other hand, there are a lot of open access resources available on the Web. This has led LIS authors to refer to more and more Web resources as part of their increased research productivity.

Currently (June 2008), according to DOAJ ^[1](2008) there are 3563 open access journals. Among them, 200 journals including Information Research are indexed in ISI (ISI, 2004). *Information Research* is a freely available, international, scholarly journal, dedicated to making accessible the results of research across a wide range of information-related disciplines. The journal is indexed by Google Scholar, INSPEC: Engineering Village, Library, Information Science & Technology Abstracts and LISA: Library and Information Science Abstracts. As mentioned above, it is also indexed for ISI's Web of Knowledge. It has also received several awards. Moreover, the journal is listed in the catalogues and directories of resources of several hundred university and college libraries around the world and in the major Internet searching tools (Information Research, 2008). Hence, the present article aims to do a detailed study on use of Web citations and citation behavior of the authors of articles published in *Information Research*. As Maharana and his colleagues (2006) declare that "citing web resources properly according to an established style is important in most of the subject fields and it is different from citing traditional resources. Apart from the style of web citations, quality, authenticity and sustainability are the issues with documents on the Web, demanding the immediate concern of the information professional"(Maharana, Nayak, & Sahu, 2006), doing studies like this seems to be important as well as necessary.

Research Objectives

The main aim of the present paper is to study and analyze Web citations of the scholarly articles published in *Information Research* during April 1995 to March 2008. However, the other objectives of the study can be summarized as follows:

- To determine distribution of articles, citations and Web citations.
- To determine distribution of articles which have Web citations.
- To find out a ratio of the citations from printed resources and those from Web resources (namely, to find out the composition of bibliographic citations).
- To find out the types of domains from where the authors have cited (i.e. to find out

the distribution of URLs by type of domain).

- To explore different file formats of the cited Web documents (i.e. to find out the distribution of URLs by type of file formats).
- To study accessibility or availability status of URLs.
- To identify error messages for inaccessible URLs.
- To calculate half-time of Web resources referred in *Information Research*.

Literature Review

Importance of using and studying Web citations has been considered since the emergence of the Internet. Therefore, several researches have been done in order to rate such relatively emergent phenomenon and its related issues. In relation to Web instability, Isfandyari-Moghaddam (2007) indicates that "apparently, the Web is dynamic. Do not expect consistency among its features" and here, Web citations (Isfandyari Moghaddam, 2007). Admittedly, such determinant feature has multiplied rate of studies in the field of internet research including Web citation analysis and related areas. Consequently, a significant body of literature exists in the area of study of the effect of Web sources on scholarly communication and the citation behavior of the authors in citing Web sources. For instance, some related studies are here explained.

Harter and Kim (1996) in an article entitled "Electronic journals and scholarly communication: a citation and reference study" done one of the first studies on availability and permanency of URLs. The major purpose of their research was to study the effects of scholarly, peer-reviewed e-journals on formal scholarly and scientific communication, as measured by cited references. Accordingly, they extracted and examined 47 unique URLs of 39 scholarly, peer-reviewed e-journals published during 1993 to 1995 (roughly the end of 1995). Finally, research findings showed that number and percentage of accessible references were 31 and 66, respectively (Harter & Kim, 1996).

Zhang (1998) investigated how much electronic sources have been used in formal scholarly communication, using a case study in the area of LIS during 1994–1996. As measured by e-resources cited, the study found that the impact of e-sources in scholarly communication in LIS is small as compared to that of print resources (Zhang, 1998).

Koehler (1999, 2002), in a longitudinal study, examined both the accessibility and content of 360 randomly chosen URLs obtained from WebCrawler over 3 years. He found that about 50% of them were still active at the end of this time and most had changed in content (Koehler, 1999 & 2002).

Germain (2000) did a research to investigate the reliability of URLs in academic citation. To do her study, 31 randomly chosen academic journal articles, containing 64 citations with URLs, were reviewed. It is worth saying that the academic journals used were from a variety of disciplines. 13 citations were from information and library science,

10 from the hard sciences, 17 from computer science, 11 from the humanities, and 13 from the social sciences. The printed journals were published between 1995 and 1997. Results of this longitudinal study found an increasing decline in the availability of URL citations. Statistically, after a three-year period, almost 50 percent of the URL citations could not be accessed and two-thirds of the journal articles contained corroded citations (Germain, 2000).

Davis and Cohen (2001) made a citation analysis of undergraduate term papers in microeconomics and revealed a significant decrease in the frequency of scholarly resources cited between 1996 and 1999. The citation analysis found that during the period under study, book citations decreased from 30% to 19%, newspaper citations increased from 7% to 19%, and Web citations increased from 9% to 21%. Furthermore, Web citations checked in 2000 revealed that only 18% of URLs cited in 1996 led to the correct Internet document (Davis & Cohen, 2001). It should be reminded that Davis (2002) in a 2000 update to the 1996-1999 citation analysis mentioned above concluded that 65% of the citations pointed directly to the cited document, up from 55% in 1999 (Davis, 2002).

Dellavalle et al (2003) have examined systematically the extent of Internet referencing and Internet reference activity in medical or scientific publications in more than 1000 articles published between 2000 and 2003 in the *New England Journal of Medicine*, *The Journal of the American Medical Association*, and *Science*. They found that Internet references accounted for 2.6% of all references (672/25548) and in articles 27 months old, 13% of Internet references were inactive (Dellavalle et al, 2003).

Casserly and Bird (2003) examined 500 internet citations randomly chosen from scholarly articles published in library and information science journals. They found that only 56.4% of those URLs were permanent, while the rest had disappeared from the original web address. Further, the study showed that more than half of the online citations contained incomplete information and the majority did not include a retrieval date. In addition, "file not found" was the most frequent error message reported by their study. They also found that close to half of the online citations they examined were initially unavailable, but increased the final result to 81.4 percent available citations by using different methods, including correcting errors in the URL, browsing the parent website or using the Google search engine (Casserly & Bird, 2003).

McCown, Chan, Nelson & Bollen (2001) explored the availability and persistence of URLs cited in articles published in *D-Lib Magazine*. For doing their research, they extracted 4387 unique URLs referenced in 453 articles published from July 1995 to August 2004. In conclusion, it was realized that approximately 28% of those URLs failed to resolve initially, and 30% failed to resolve at the last check. A majority of the unresolved URLs were due to 404 (page not found) and 500 (internal server error) errors. Moreover, based on the data collected, they found the half-life of a URL referenced in a D-Lib Magazine article

is approximately 10 years. It was also found that URLs were more likely to be unavailable if they pointed to resources in the .net, .edu or country-specific top-level domain, used non-standard ports (i.e., not port 80), or pointed to resources with uncommon or deprecated extensions (e.g., .shtml, .ps, .txt) (McCown, Chan, Nelson & Bollen, 2001).

In a study, Maharana and his colleagues (2006) analyzed 292 Web citations spread over 95 scholarly papers published in the proceedings of the National Conference of the Society for Information Science, India (SIS-2005). All the 292 web citations were scanned and data relating to types of web domains, file formats, styles of citations, etc., were collected through a structured check list. Ultimately, the study revealed that 292 (34.88%) out of 837 were web citations, proving a significant correlation between the use of Internet resources and research productivity of LIS professionals in India. The highest number of web citations (35.6%) was from .edu/.ac type domains. Furthermore, most of the web resources (46.9%) cited in the study were html files (Maharana, Nayak & Sahu, 2006).

Wren, Johnson, Crockett, Heilig, Schilling & Dellavalle (2006) explored URL decay in dermatology journals. To do this, they considered URLs in articles published between January 1, 1999, and September 30, 2004, in the 3 dermatology journals with the highest scientific impact. The percentage of articles containing at least 1 URL increased from 2.3% in 1999 to 13.5% in 2004. Of the 1113 URLs, 81.7% were available (decreasing with time since publication from 89.1% of 2004 URLs to 65.4% of 1999 URLs) ($P < .001$) (Wren, Johnson, Crockett, Heilig, Schilling & Dellavalle, 2006).

Dimitrova and Bugeja (2007), in an exploratory study, examined the use of online citations, focusing on five leading journals in journalism and communication. In operation, they analyzed 1126 URL reference addresses in citations of articles published between 2000 and 2003. The results showed that only 61% of the online citations remain accessible in 2004 and 39% do not. The content analysis also demonstrated that .org and .gov are the most stable domains (Dimitrova & Bugeja, 2007).

In addition to aforementioned works, Falagas and colleagues (2007) comparatively examined the frequency and the specific problems encountered in accessing Internet references in two leading medical journals during the last 3 years. In their study, two investigators independently reviewed all publications in the issues of the *New England Journal of Medicine* and *The Lancet* during October 2005 to March 2006, November 2004 to January 2005, and November 2003 to January 2004. They calculated the total number of references and the subset referred to an Internet source of each article. Then, they visited the electronic sources to identify the Internet references and noted the problems of accessibility, if any. When they failed to directly access the reference in the electronic address provided by the authors, they visited the referred website; if this was also inadequate, they performed Google searches to retrieve the missing reference(s). It was

finally found that 465/18,850 (2.5%) and 952/24,630 (3.9%) of the reviewed references in the *New England Journal of Medicine* and *The Lancet*, respectively, referred to Internet sources; from these they could not access 68/465 (14.6%) and 170/952 (17.9%) in the two journals, respectively. Additionally, there were increasing proportions of lost Internet references as they age. Searching into the website referred by the authors of the reviewed articles could not provide the missing information in a considerable proportion (62.2%). However, the use of an Internet search engine (Google) helped them to identify references in other websites, reducing the proportion of missing Internet references to 17/465 (3.7%) and 17/952 (1.8%) for the two journals, respectively. It is notable that the response "page not found" was commonly encountered when they tried to access Internet references in publications of the studied journals (Falagas, Karveli & Tritsaroli, 2007).

So, building this research on previous works mentioned above and using a single case study approach, this article aims to explore availability and half-life of URLs cited in articles published by *Information Research* journal.

Research Method

At first, we extracted all issues of *Information Research* from April 1995 to March 2008 and calculated number of all citations both printed citations and Web ones. It should be noted that only articles which had references have been studied. Accordingly, editorials, reports, reviews and so on which had no references have been neglected. Afterwards, we checked availability of individual cited URLs. When we could not access directly e-address (link) inserted in any article by author(s), we tried to view referred website. If this attempt seemed to be insufficient, search engine "Google" was employed to access the missing reference(s). Finally, collected data were analyzed, and needed tables as well as figures were drawn using Excel software in accordance with research objectives.

Research Findings

In the present study, the Web resources referred to by authors in *Information Research* articles in the references section of their papers have been studied. A total number of 10242 citations, as obtained from the bibliographies of 339 papers, have been analyzed and the necessary interpretations were made. Here, in line with research objectives mentioned earlier, research findings are provided as below:

Distribution of articles, citations and Web citations

As shown in Table 1, totally, 339 articles have been published in *Information Research* during years 1995-2008. It is important to note that year 2004 with 46 articles and year

1995 (the first year of publication of *Information Research*) with 6 articles have the most and the least rate of articles, respectively. According to Table 1, there were total number of 10242 citations in 339 articles, and average "30.21 citations" was calculated per paper. Additionally, among all citations (10242), there are 1761 Web citations with the average "5.19 Web citations" per paper. Comparing these results with findings of Casserly and Bird (2003) reveals that average of citations (30.21) and Web citations (5.19) per paper in *Information Research* is more than average of citations (25) and Web citations (2.5) per paper in the journals studied by them.

Table 1
Distribution of articles, citations, and Web citation in Information Research

Year	Articles evaluated	Total citations	Average citation per paper	Total Web citations	Average Web citation per paper
1995	6	95	15.83	6	1
1996	14	189	13.5	5	0.35
1997	21	287	13.66	32	1.52
1998	14	290	20.71	17	1.21
1999	14	322	23	33	2.35
2000	23	679	29.52	99	4.30
2001	31	954	30.77	85	2.74
2002	23	597	25.95	147	6.39
2003	21	995	47.38	213	10.14
2004	46	1190	25.86	230	5
2005	32	1213	37.90	134	4.18
2006	43	1571	36.53	314	7.30
2007	42	1424	33.90	365	8.69
2008	9	436	48.44	81	9
All years	339	10242	30.21	1761	5.19

Distribution of articles with Web citations

As can be seen in Table 2, of 339 published articles 224 ones have Web citations. In a word, 66% of articles, at least, have one Web citation.

Table 2

Distribution of articles with Web citation in Information Research

Year	Articles evaluated	Articles with Web citation	Percentage of articles with Web citation
1995	6	1	17%
1996	14	3	21%
1997	21	5	24%
1998	14	3	21%
1999	14	3	21%
2000	23	13	57%
2001	31	15	48%
2002	23	20	87%
2003	21	19	90%
2004	46	40	87%
2005	32	23	72%
2006	43	40	93%
2007	42	31	73%
2008	9	8	89%
All years	339	224	66%

Based on findings included in Table 2, it can be said that LIS authors have been increasingly making use of Web-based resources in their studies. Also, it is of importance that these results are considerably different from Wren and colleagues' (2006) because they reported that 13.5% of dermatology articles have Web citations, while, as illustrated in Figure 1, *Information Research* articles have growingly used Web citations from 17% in 1995 to 89% in 2008.

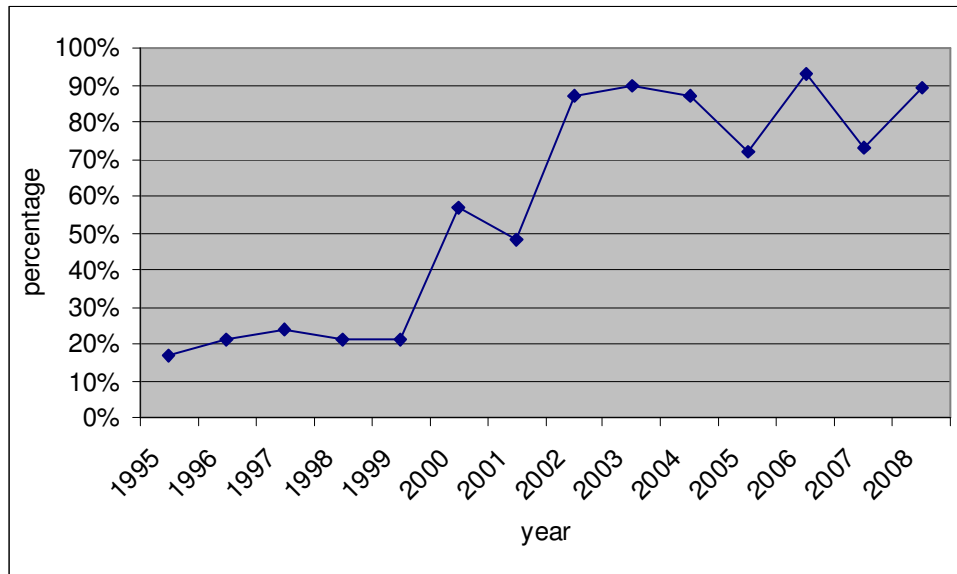


Figure 1. Percentage of articles with Web citation in *Information Research*.

Composition of bibliographic citations

According to Figure 2, of total citations (10242), citing to Web and printed resources in *Information Research* articles is 1761 (17%) and 8481 (83%), respectively. It demonstrates that paying attention to Web citations in *Information Research* articles is considerable compared to Web citations used in medical journals (~3%), as reported by Falagas and colleagues (2007).

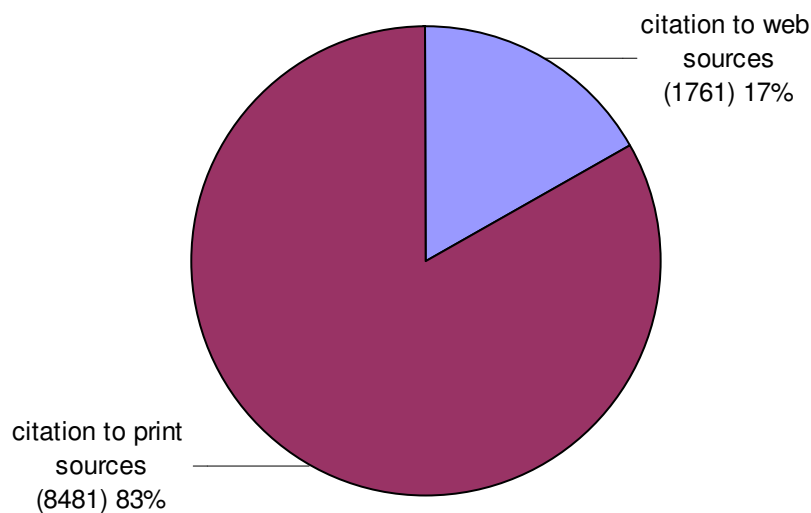


Figure 2. Composition of bibliographic citations in *Information Research*.

Distribution of URLs by type of domain

As said by Maharana and his colleagues (2006), the URL is the address of the location of a digital document on the Web. A URL essentially has four parts: protocol, domain, directory and file. A domain name is a way to identify and locate computers connected to

the Internet. No two organizations can have the same domain name. A domain name always contains two or more components separated by periods, which are called "dots". Some examples of domain names are: ibm.com, nasa.gov, utexas.edu and tcs.co.in. A domain name can often tell the user if it is a government site, an academic site or a commercial site. Some common top-level domain name endings are:

- .com or .co: a commercial organization.
- .edu or .ac: an educational organization.
- .gov: an official government site.
- .org: mostly non-profit organizations.
- .net: traditionally it was for network organizations, but now can be used by anyone.

Country codes: A two-letter abbreviation for a particular country. For example, "ir" for Iran, "UK" for United Kingdom and "Fr" for France, etc.

In this research, six different types of domain have been taken into consideration. They are .org, .edu/.ac, .com, .gov, .net and .uk while those domains not falling into any of these categories fall into the "other" category.

As Figure 3 shows, the domains of the cited URLs mostly include the .com/.co and .org types. Accordingly, of 1761 Web citations, the highest number of domains, i.e. 463 cases are of .org type and 319 ones are of .com/.co type. This reveals that the data sources of most of the Web citations in the present study are websites of various professional institutions or societies and commercial organizations, and the like. Also, 206 domains are of .edu/.ac, 198 domains (.net), 186 ones (.uk), 79 citations (.gov) and 310 cases belong to domains other than the above types, categorized in the form of "others". It should be considered that the stability and persistence of .net and .org domains are better than that of .edu, .gov, .uk, and .com sites. It is notable that these findings are compatible with McCown and colleagues (2001) in which .org domains had more sustainability than .edu domains.

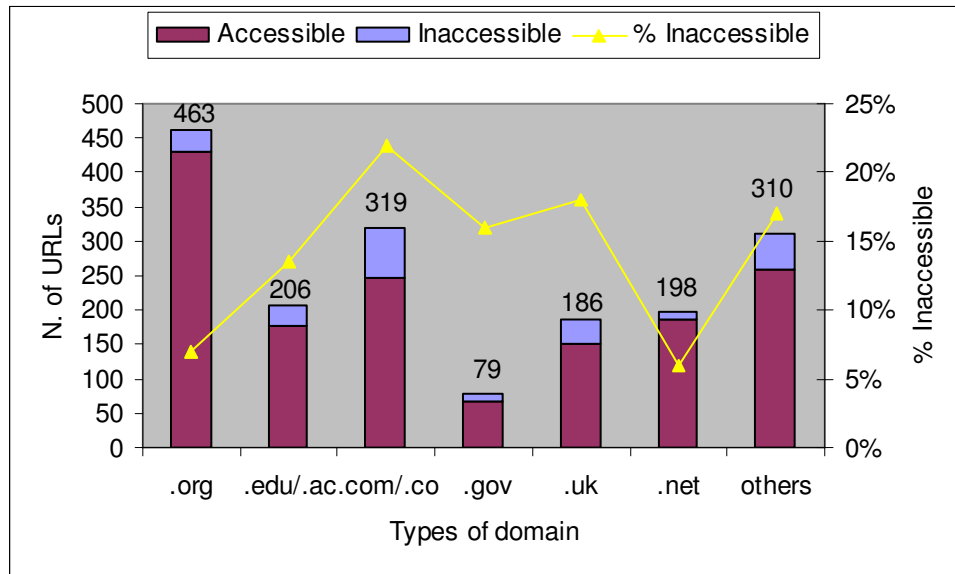


Figure 3. Distribution of URLs by type of domain.

Distribution of URLs by type of file formats

Similar to previous studies, McCown, Chan, Nelson & Bollen (2001) and Maharana, Nayak & Sahu (2006), the URLs have been categorized into seven different file formats as follows:

- Slash files (/): URLs which end by / sign, for example, <http://foo.edu/>.
- HTML (hyper text markup language): Web documents created in HTML scripting language.
- PDF (portable document format): the file format for documents created using Adobe Acrobat.
- PPT: PowerPoint presentations.
- DOC: documents created using MS-Word.
- RTF (rich text formats): a text file format that includes formatting features, such as bold, italic, and underlined text.
- Others.

The data as illustrated in Figure 4 indicate that the greatest numbers of cited web resources are HTML/HTM files. Out of 1761 Web citations, 884 cases are HTML files, followed by 371 PDF files, 225 Slash files, 16 DOC files, 4 RTF files and 4 PPT files. Some other file formats, 257 ones, which did not match these six categories, were included in the "other" category.

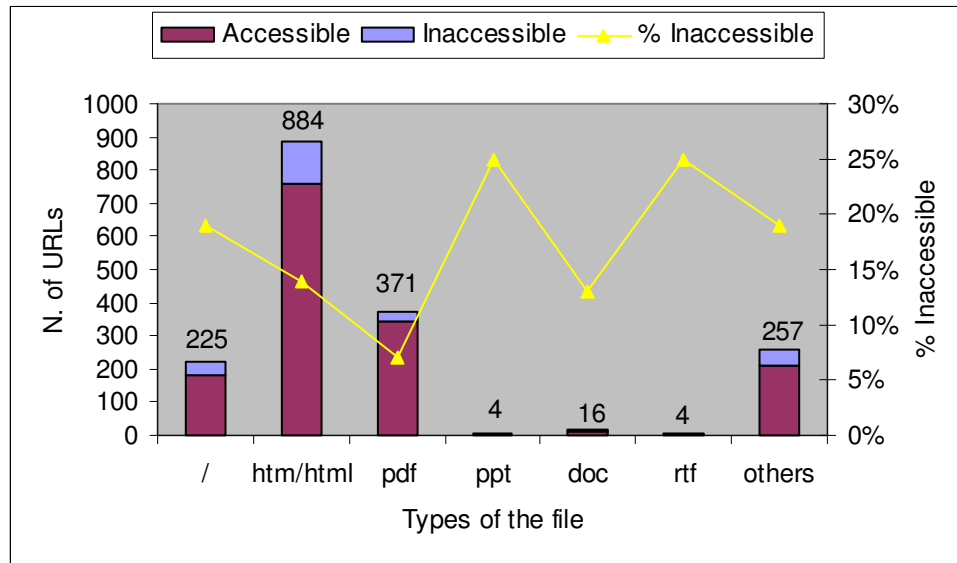


Figure 4. Distribution of URLs by type of file formats.

Our findings relating to this section are supported by findings of McCown, Chan, Nelson and Bollen (2001) and Maharana, Nayak & Sahu (2006). Their research reported that most of cited Web resources contain HTML/HTM files. Additionally, as reflected in Figure 4, PDF files are the most stable files. In McCown, Chan, Nelson, and Bollen (2001), the most stable files were also PDF files.

Availability of cited URLs at first check

Figure 5 demonstrates that of 1761 cited URLs, 1290 (73%) were accessible, while only 471 (27%) were inaccessible. In other words, 1290 URLs directly led to referred resources. Yet, in order to access 471 URLs we encountered errors. Consequently, as mentioned earlier, when we failed to directly access the reference in the electronic address provided by the authors, the referred website was visited. If this attempt was also insufficient, search engine "Google" was employed to access the missing reference(s).

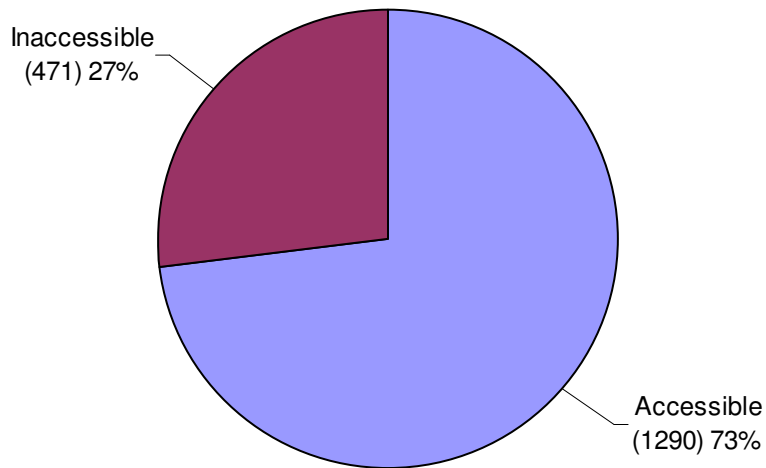


Figure 5. Availability of URLs at first check.

The status of the availability of cited URLs after searching through Google has been illustrated in Figure 6.

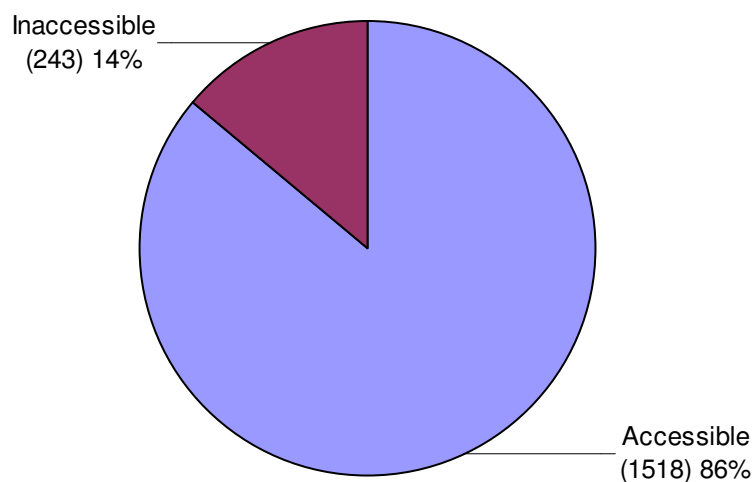


Figure 6. Availability of URLs after searching at Google.

As reflected in Figure 6, after searching missing URLs via Google, accessible URLs increased from 1290 (73%) to 1518 ones (86%). Conversely, inaccessible URLs decreased from 471 (27%) to 243 cases (14%).

Status of accessible URLs

Figure 7 demonstrates how accessible URLs were found. Accordingly, of total number of accessible URLs, 835 cases (55%) were found at cited URL, 420 cases (28%) were found at Internet archive, 220 ones (14%) were accessed using Google search engine, 35 URLs (2%) were found at another URL other than cited URL, and 8 cases (1%) were accessed through searching missing URL in the Internet.

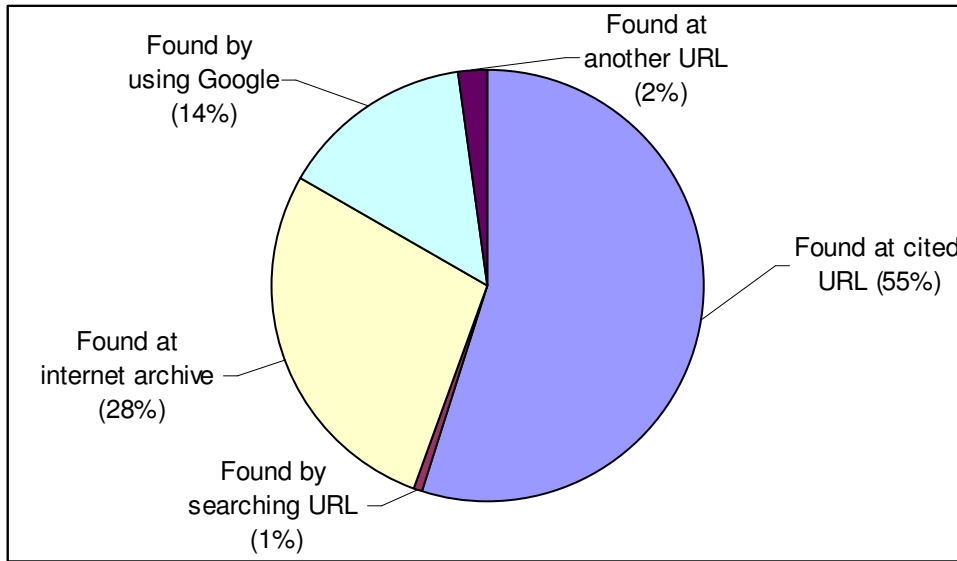


Figure 7. Accessible URLs

Error messages at inaccessible URLs

In relation to 243 inaccessible URLs, error messages have been recorded which can be seen in Figure 8.

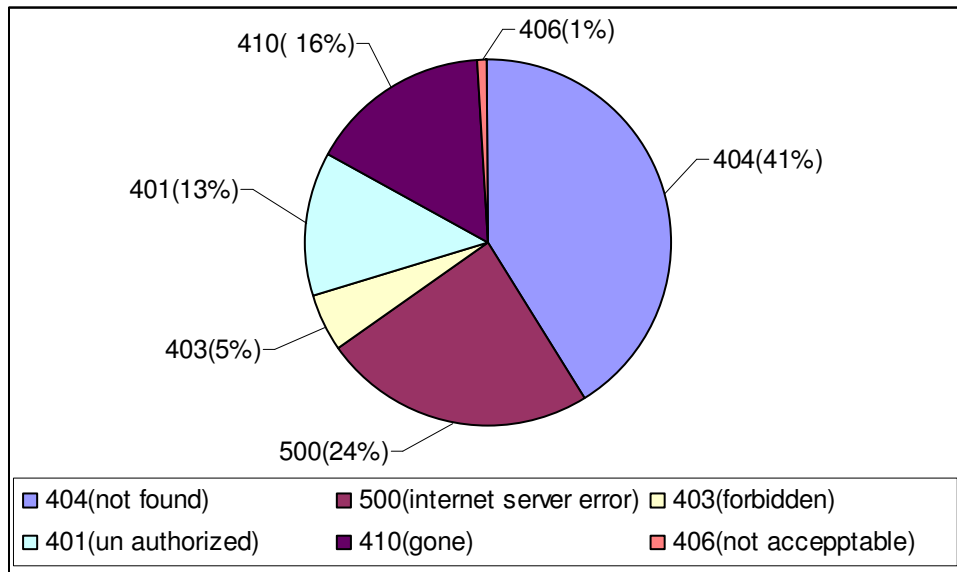


Figure 8. Error messages at inaccessible URLs.

As shown in Figure 8, errors 404 (not found) with 41% (100 messages) and 500 (server error) with 24% (58 messages) are the most recorded error messages, respectively. Moreover, error messages 410 (gone) with 16% (39 cases), 401 (unauthorized) with 13% (31 ones), 403 (forbidden) with 5% (13 messages), and 406 (not acceptable) with 1% (2 messages) were recorded. It should be reminded that these findings are compatible with McCown, Chan, Nelson & Bollen (2001) in which decay and instability of most of Web

citation were because of errors 404 (not found) and 500 (server error).

Half-life of Web citations in Information Research

In order to estimate half-life of Web resources cited in *Information Research* articles, the procedure used in previous research by Koehler (1999), Tyler and McNeil (2005), and Dimitrova and Bugeja (2007) has been employed. By definition, "half-life" is the time required for half of all online [Web] citations in a journal to disintegrate. This amount of time may differ for different disciplines or different years (Koehler, 1999). The following formula was used to calculate the half-life of online citations for each journal year:

$w(t) = w(0)e^{at}$, where $W(0)$ is the number of working online citations at the time of publication, $W(t)$ is the number of working online citations at some later time t , and a is a constant that can be calculated from the available data. Now, the half-life of Internet citations, t_h , is calculated as follows:

$t_h = [t \ln(0.5)] / [\ln w(t) - \ln w(0)]$, where t_h is the estimated number of years it takes for 50% of the published Internet citations to stop working. Using the second formula, we calculated the half-life for *Information Research*. Results of this estimation are included in Table 3.

Table 3
Half-life of Web citations in Information Research

Year	Half-life
1995	14
1996	17.63
1997	10.06
1998	10.11
1999	24.96
2000	8.13
2001	7.86
2002	9.05
2003	6.65
2004	8.92
2005	10.93
2006	11.48
2007	29.06
2008	27.72
Average	14.04

As can be observed in Table 3, the average half-life for cited Internet resources is estimated to be 14.04 years. This means that it will take about fourteen years for half of the Internet citations in *Information Research* to vanish. Also, based on Table 3, year 2007 with the average half-life, 29.06, and year 2003 with the average half-life, 6.65, had the highest and the least average half-life, respectively.

Conclusion

Today, the Internet has become a valuable, perhaps indispensable resource in conducting scientific research (Lawrence & Giles, 1999), not just because of the added convenience of rapid information retrieval and sharing, but because it also provides a means of making resources available that the printed media simply cannot. Therefore, even though the authors may appreciate the risk of future inaccessibility of Internet references, they cannot easily avoid their use in their publications (Falagas, Karveli & Tritsaroli, 2007). From what the present research showed us, it can be said that, as declared by Dimitrova and Bugeja (2007), the Internet may prove to be an inhospitable medium, especially for web-based research, because Web citations are speedily as well as constantly fading away. In spite of this, it should be accepted that the use of Internet for identifying valuable and timely information has become inevitable for most scientists as well as the public with access to the World Wide Web, since scientific and other work is created and added in digital format on the Internet every day. Such idea is reinforced by Dimitrova and Bugeja (2007) who declared as following:

"Internet research is vital to scholarship because the medium serves as a convenient electronic warehouse of data accessible at all hours and in great quantities, thereby increasing the scope and breadth of scholarship".

In order to increase the rate of availability of URLs, Wren, Johnson, Crockett, Heilig, Schilling & Dellavalle (2006) suggested that publishers, editors, and authors should work together through:

1. Requiring authors to retain digital backup or printed copies of cited Internet-only information to facilitate content recovery should a URL become unavailable.
2. Advocating the inclusion of referenced Internet content in an online archive.
3. Checking URLs systematically before publication to minimize unavailability due to spelling errors or misprints.

In addition to considering above recommendations and also, using domains and files which are more stable and persistent, we indicate that the best solution to prevent decay or disappearance of Web citations and diminish URLs decay is to make use of WebCite[®]-enhanced reference. WebCite[®], a member of the International Internet Preservation Consortium, is an on-demand archiving system for Web references (cited Web pages and websites, or other kinds of Internet-accessible digital objects), which can be used by

authors, editors, and publishers of scholarly papers and books, to ensure that cited Web material will remain available to readers in the future. A WebCite[®]-enhanced reference is a reference which contains – in addition to the original live URL (which can and probably will disappear in the future, or its content may change) – a link to an archived copy of the material, exactly as the citing author saw it when he accessed the cited material (WebCite, 2008a). Since its official launch in October 2005, more than 100 journals are already using WebCite on a routine basis (WebCite, 2008b). Based on WebCite (2008b), Information Research is observed in the list of all active members of the WebCite[®] Consortium.

Results of the availability of *Information Research* URLs studied in the present research revealed that about 99% of URLs uploaded in WebCite[®], i.e. URLs which made use of WebCite[®]-enhanced reference, were accessible. It is important to note that although *Information Research* has been member of the WebCite[®] Consortium since about 2006, we witnessed that most of URLs of years 2006, 2007 and 2008 were not uploaded. It is perhaps because *Information Research* has not apparently indicated this issue in section "Instructions for Authors" and has not necessitated authors to utilize WebCite[®] for all Web citations. Thus, it is suggested that all scholarly journals particularly *Information Research* as an international well-known journal in Library and Information Science area call for using WebCite[®]-enhanced reference and oblige authors to utilize WebCitation.org for all of citations referred in their articles. Hereby, it is also reminded that International Journal of Information Science and Management (IJISM), formerly called International Journal of Information Science and Technology (IJIST), value for WebCite[®]-enhanced reference. On the other hand, as for further research, doing a similar study which examines availability and half-life of URLs cited in articles published by IJISM from 2003-2009 is of importance.

Acknowledgements

Special thanks to Dr. Shah-Shojaee for his useful comments.

Endnote

1. Directory of Open Access Journals

References

- Casserly, M. F. & Bird, J. E. (2003). Web citation availability: Analysis and implication for scholarship. *College & Research libraries*, 64 (4), 300-317.
- Davis, P. M. & Cohen, S. A. (2001). The Effect of the Web on undergraduate citation behavior 1996-1999. *Journal of the American Society for information Science and technology*, 52 (4), 309-314.

- Davis, P. M. (2002). The Effect of the Web on Undergraduate Citation Behavior: A 2000 Update. *College and Research Libraries*, 63 (1), 53–60.
- Dellavalle, R. P., Drake, A., Graber, M., Heilig, L., Hester, E., Kuntzman, J. & Schilling, L. (2003). Going, going, gone: Lost internet references. *Science*, 302 (5646), 787-88.
- Dimitrova, D. V. & Bugeja, M. (2007). The half-life of internet references cited in communication journals. *New Media & Society*, 9 (9), 811–826.
- DOAJ (2008), Directory of Open Access Journals. Retrieved June 3, 2008, from <http://www.doaj.org> .
- Falagas, M. E., Karveli, E.A. & Tritsaroli, V. I. (2007), The risk of using the internet as reference resource: A comparative study. *International Journal of Medical Informatics*, 1–7.
- Germain, C. A. (2000). URLs: Uniform resource locators or unreliable resource locators. *College and Research Libraries*, 61 (4), 359–65.
- Harter, S. P. & Kim, H. J. (1996). Electronic journals and scholarly communication: A citation and reference study. *Information Research*, 2 (1), paper 9. Retrieved June 25, 2008, from <http://InformationR.net/ir/2-1/paper9a.html>.
- Information Research (2008). Information Research: an international electronic journal. Retrieved June 25, 2008, from <http://informationr.net/ir/about.html>.
- Isfandyari Moghaddam, A. (2007). Web metasearch engines: A comparative study on search capabilities using an evaluation check-list. *Online Information Review*, 31 (3), 300-309.
- ISI (2004). The Impact of Open Access Journals: A Citation Study from Thomson ISI. Retrieved June 25, 2008, from <http://scientific.thomsonreuters.com/media/presentrep/acropdf/impact-oa-journals.pdf> .
- Koehler, W. (1999). An analysis of web page and website constancy and permanence. *Journal of the American Society of information Science and Technology*, 50 (2), 162-180.
- Koehler, W. (2002). Web page change and persistence—a four-year longitudinal study. *Journal of the American Society of information Science and Technology*, 53, 162–171.
- Lawrence, S. & Giles, C. L. (1999). Accessibility of information on the web. *Nature*, 400, 107–109.
- Maharana, B., Nayak, k. & Sahu, N. K. (2006). Scholarly use of web resources in LIS research: A citation analysis. *Library Review*, 55 (9), 598-607.
- McCown, F., Chan, S., Nelson, L. M., Bollen, J. (2001). The availability and persistence of web references in D-Lib magazine. Retrieved June 25, 2008, from <http://www.iwaw.net/05/papers/iwaw05-mccown1.pdf>.

Tyler, D.C. & McNeil, B. (2003). Librarians and link rot: A comparative analysis with some methodological considerations. *Portal: Libraries and the Academy*, 3 (4), 615–632.

WebCite (2008a),.Home. Retrieved June 25, 2008, from <http://www.Webcitation.org/index>.

WebCite (2008b). WebCite Consortium Members. Retrieved June 25, 2008, from <http://www.Webcitation.org/members>.

Wren, J. D., Johnson, K. R., Crockett, D. M, Heilig, L. F., Schilling, L. M., Dellavalle, R.P. (2006). Uniform resource locator decay in dermatology journals. *Arch Dermatol*, 142, 1147-1152.

Zhang, Y. (1998). The impact of Internet based electronic resources on formal scholarly communication in the area of library and information science: A citation analysis. *Journal of Information Science*, 24 (4), 241–254.

