

## **An Analysis of the Availability and Persistence of Web Citations in Iranian LIS Journals**

**A. H. Mardani, M.S.**

Tehran University of Medical Sciences, Iran

Corresponding author:

mardani3@gmail.com

**M. Sangari, M.S.**

Payam-e-Noor University

Mashhad Branch, Iran

email: msangary@gmail.com

### **Abstract**

To discover the current situation and characteristics of web citations accessibility, the present study examined the accessibility of 4,253 web citations in six key Iranian LIS journals published from 2006 to 2010. The proportion percentage of web citations increased from 11% in 2006 to 30% in 2010. The most widely cited top level domains in URLs include the .edu and .org with respectively 37% and 23%. This study provides further evidence that organizations websites have become increasingly vulnerable to URL decay. The results show that only 3467 web citations remain accessible in 2011, of which 71% allowed easy and long-term access to the authors' information intended in URLs. Long time inaccessibility to the authors' intended information was shown to be mostly from URLs that returned the 404 error and also the URLs that had gone through information update. An about 4 year half-life was estimated for Iran's LIS Publications. Ultimately, the results suggest that the decay of URLs is a grave problem in the publication of Iran's LIS researchers and cannot be overlooked. These authors need to gain the necessary knowledge about using web citations as major sources of information for their publications.

**Keywords:** Web Citation, URL, Iran, Internet Resources Half-life, LIS journals.

### **Introduction**

The Internet is one of the most important and complex innovations in human history, the largest and most complete tool for information exchange ever made available to the global population. Since the quasi-miraculous emergence of the web in 1990s, there has been a continuous increase in the volume of scholarly resources in electronic forms 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, 2006).

Web resources play a prominent role as a medium of information in helping Iranian

researchers with their studies and it has been supported in numerous theoretical and empirical studies (Sa'adat, 2008; Roshandel, Rezailashkajani, Ansari, & Zali, 2005). Web has gradually become an information source in the research papers and publications in Iran and has continued to work its way through different scientific fields but the fleeting nature of information on the internet and rapid changes in web technologies have concerned both LIS authors and publishers. The constantly changing environment of the internet does not provide any guarantee of permanence. Anyway, mobility and inaccessibility of web-based sources are still serious reasons that may affect scholars' tendency to use them as legitimate media of formal scholarly communication (Spinellis, 2003).

In this study, we investigate how web citations are used in articles published in Iranian journals during 2006 to 2010 and analyze the results. This study assessed the overall rate of web citations and the frequency of inaccessible URLs in LIS journals. Although cross-disciplinary web citations allow us to compare differences among disciplines regarding the acceptance to web resources in scholarly publications, concentrating on one discipline can provide a deep understanding of uses of web resources in scholarly publications.

### Review of Literature

Shortly after the web emerged, information scientists started to monitor the impact of web resources on the scholarly literature and the citation behavior of authors during the process of research. Harter & Kim (1996) studied citations in 279 articles published in peer-reviewed LIS e-journals to measure the extent to which authors cite e-journals and other online sources. Harter & Kim found that while e-journals accounted for 0.2 percent of the references cited, citations of various online sources as a whole consisted of 1.9 percent of the total references. In a similar citation study, Zhang (2001) investigated the use of internet-based electronic resources by LIS researchers for scholarly contributions. Zhang discovered that the percentage of electronic citations in print journals of LIS increased from 0.2 to 5.2 percent during the period of 1991 and 1998, whereas the percentage of articles containing electronic citations showed a dramatic increase from 1.8 to 33.9 percent.

Markwell & Brooks (2002) found that some 14% of all URLs had ceased to function or changed their content in the first 14 months of the study. The authors found that the top-level domain that was most reliable and less likely to disappear when compared to other top-level domains were those associated with government websites. Conversely, the .com domain exhibited the highest degree of instability with almost 50% of these types of URLs not being accessible after 24 months.

Casserly & Bird (2003) conducted a study of LIS journal articles published between 1999 and 2000 and found that among the 35,689 citations included in 1,425 articles, a total of 3,582 or 10% of the references were originated from the web.

In a study of web citations in high-impact oncology journals, Hester (2004)

documented an increase in the number of articles with references to websites, from 9% in 2001 to 16% in 2003. The proportion of websites no longer available also increased. Five months after publication, 9.5% of the internet addresses were inactive, and at 29 months after publication, 33% of the web citations were no longer available.

Maharana & Nayak (2006) investigated 1126 URL reference addresses in citations of articles published between 2000 and 2003. The results showed that only 61 percent of the web citations remained accessible in 2004 and 39 percent do not. The content analysis also shows that .org and .gov are the most stable domains. Error messages for 'dead' URL addresses are explored.

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

Wagner, Gebremichael, Taylor & Soltys (2009) examined the accessibility and decay of web citations in the Healthcare Management journals during 2002 to 2004. To do this, they extracted 2011 unique URLs from 5 accredited journals in the field of Healthcare Management. The accessibility and decay of URLs suggested that 50.7% of URLs were accessible and 49.3% returned error messages and thus were inaccessible. The results also revealed that the .edu and .net domains were the most stable domains with accessibility rates of 68.4% and 61.5%, respectively.

Isfandyari Moghaddam & Saberi (2010) examined the accessibility and half-life of cited URLs in the published papers in the Information Research Journal. 66% of the papers had web citations and the rate of using web citations as references had increased over the past four years. The .org and .net domains had the most stability and 73% of the URLs were accessible. It is also notable that using Google and searching the missing URLs, the accessibility had increased up to 86%.

### **Research Questions**

The purpose of this study is to add to the knowledge about the changing landscape of scholarly communications by examining citations to web resources included in articles published in the LIS. Specifically, this study addresses the following questions:

- What proportion of the web citations has been used by authors in their articles?
- How much is the half-life of the web resources cited in the articles?
- What is the overall rate of URL decay for web citations in LIS articles?
- What top-level domains constitute the source of the web-located citations?
- What is the proportion of the top-level domains in inaccessible citations?

### Research Methodology

The present research investigates the available web citations in LIS articles by Iranian researchers published in journals during 2006 to 2010. Here, web citation means a reference to any resource that has a unique web URL presence, regardless of whether there is a print equivalent to that web resource. The articles pertain to six prominent LIS journals published in Iran including: International Journal of Information Science and Management (IJISM), Faslname- ye- Ketab, Journal of Information Science & Technology, Research on Information Science & Public Libraries, Library and Information Science, Informology. These are leading LIS journals published in Iran and researchers have published their recent findings in them. In addition, the mentioned journals have regularly been published during 2006 to 2010 and without interruption.

Access to the articles of the journals has been provided via their websites. In the present research, the web resources referred to by authors in the reference sections of their articles have been studied. It should be noted that only articles which had references have been studied. Accordingly, editorials, reports, reviews and etc. which had no references have been neglected.

A total number of 18801 citations, as obtained from the bibliographies of 1032 articles presented, have been analyzed and the necessary interpretations were made. After lengthy manual verification, total number of the research articles published during the five years, number of articles with web resources, number of total citations and number of citations to web resources were recorded into separate Excel files. The URL of each web resource was also noted for determining the nature of the web resource and its availability in 2011. Two research assistants checked the availability of the web citations and recorded the information in an Excel file.

In order to fulfill the primary purpose of this paper which was to offer some information about the status of accessibility of web citations, we attempted to analyze the accessibility of every single URL used in all the web citations in the LIS articles. Throughout a 2 month period (January and February 2011) we attempted to access all URLs cited in articles during our study years. Therefore, based on the Thorp & Brown's (2007) classification we categorized web citations into 9 groups and for the sake of a better interpretation of data, we organized the 9 groups into two main categories. The first group consisted of web citations that were readily accessible and contained information that the author intended to cite. The second consisted of web citations that failed to contain readily accessible information that the author intended to cite. The research assistants faced many challenges in trying to determine whether content at a given URL matched that viewed by the author of the article in which it was cited. For citations that included full bibliographic information and the date the author viewed the cited content, the research assistants were able to determine with some certainty whether the content of the web page matched the

cited information. We were concerned that the dynamic nature of the internet would give the “no site found” message or “temporarily unavailable” message erroneously. To avoid misclassification when these messages appeared, we retried each site at least 24 hours after the failed attempt.

### Results

As it can be seen from the Table 1, the web citations exhibit an exponential increase from 2006 to 2010. Almost 11% (475 URLs) of the web citations was from articles published in 2006, followed by 14% (629 URLs) in 2007 and 18% (758 URLs) in 2008 and 27% (1137 URLs) in 2009 and 30% (1254 URLs) in 2010. The table shows that web citations increased from 2006 to 2010, consistent with the literature reporting the increasing popularity of web citations.

It was found that some articles included many web citations, in some cases, above 50% of the total number of citations. As a result, one can assert that such researches are fundamentally based on web information. Of course, the number of web citations varied among articles. The highest number of web citations in an individual article was 19, with some authors not citing any web-located resources. An average number of 4 was estimated for the number of web citations per article for all the studied years. Collectively, there were 824 articles containing at least one web resource, thus 80 percent of the total number of articles contained web resources. The findings concur with the findings of other researchers that the number of web resources in journals has grown over time, even though scholars appear to make limited use of web publications.

Table 1

*The Growth of Web Citations in LIS Articles (2006-2010)*

Year	Articles Evaluated	Total Citations	URLs	Average URL per Articles	Percentage Year
2006	179	2577	475	2.65	11
2007	223	3861	629	2.82	14
2008	222	3439	758	3.41	18
2009	214	4761	1137	5.31	27
2010	194	4163	1254	6.46	30
Total	1032	18801	4253	4.12	100

Despite the widespread utilization of web resources, and as suggested in the findings of the previous section of the study, many researchers have come to the conclusion that a significant number of URLs used in the web citations would become inaccessible in the years after publication. Anyhow, this matter of concern urged us to determine the half-life

of URLs in our study using the URLs half-life scale.

In order to estimate half-life of web resources cited in LIS articles, the procedure used in previous research by Koehler (1999), Tyler & McNeil (2003), and Dimitrova & 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 web citations for each journal year:

$$W(t) = W(0)e^{at}$$

Where  $W(0)$  is the number of working web citations at the time of publication,  $W(t)$  is the number of working web citations at some later time  $t$ , and  $a$  is a constant that can be calculated from the available data. Now, the half-life of web citations,  $t_h$ , is calculated as follows:

$$t_h = [t \int n(0.5)] / [\int nW(t) - \int nW(0)]$$

Here,  $t_h$  is the estimated number of years it takes for 50% of the published web citations to stop working. Using the second formula, we calculated the half-life for each of the six journals and each of the five publication years that we analyzed. The results are shown in Table 2.

Based on the five-year period and the journals examined in this study, the average half-life for LIS web citations is estimated to be 3.63 years. This means that it will take about four years for half of the web citations to vanish. Based on the information presented in Table 2, editors can estimate the projected attrition rate for following years and note whether the half-life increases or decreases, using 3.63 as a benchmark number. Certainly, the decay period for web resources is different in various disciplines, but the estimated half-life in some previous studies ranges between 1.5 to 5 years (Dimitrova & Bugeja, 2007; Sellitto, 2004) and is above 10 years in some others (Isfandyari Moghaddam & Saberi, 2010). All of these findings show that the decay of web citations in the study population could be considered a grave matter that would require further research.

Table 2

*Half- life of Web Citations in LIS Journals*

<b>Journal / Year</b>	<b>2006</b>	<b>2007</b>	<b>2008</b>	<b>2009</b>	<b>2010</b>	<b>Average</b>
IJISM	4.28	3.26	7.05	2.92	3.35	4.17
Fasname- ye Ketab	2.83	1.18	5.64	3.31	2.04	3.0
Library and Information Science	1.18	4.27	1.03	3.61	2.15	2.44
Journal of Information Science &Technology	6.24	4.18	3.55	5.0	8.07	5.40
Research on Information Science & Public Libraries	2.09	1.22	4.67	3.46	4.33	3.15
Informology	4.01	3.47	3.05	5.72	2.06	3.66
All Journals	3.44	2.93	4.16	4.0	3.66	3.63

Not surprisingly, the discontinuity of web resources is due to the dynamic nature of web pages that leads to changed content or missing information. The poor stability and persistence of web citations may be one of the factors that discourage scholars in using the resources available on the internet for scientific publication.

A general review of the data from Table 3 shows that 3467 web citations (82% of the total citations) ultimately became accessible and 786 citations (18% of the total citations). Actually, those that returned “No site found” and “Temporarily Unavailable” messages became ultimately inaccessible. As reported in other studies, most internet users encounter the famous 404 error of “page was not found” (Dimitrova & Bugeja, 2009; Falagas, Karveli & Tritsaroli, 2007).

A more careful look at the groups of citations shows that 2999 web citations (71% of the total citations) provided easy and long term access to the authors intended information and the number of citations in this group increased with the number of web citations in papers during this study and had a 5.3 growth rate. In this section, we used the bibliographic information provided in the web citations of the articles such as titles, authors’ names and publication dates to determine the constancy of the selected information.

When comparing the categories of web citations, we found that the number of web citations that no longer accessed the authors intended information decreased over time; 43 percent of web citations in 2006 to 8 percent of web citations in 2010 no longer allowed access to the authors intended information. To better describe how the web citations were inaccessible, we compared the 6 categories of web citations that did not readily link to the authors intended information. During the 5 years, 701 web citations (55%) could not locate a website and hence returned with the 404 error. 271 (21%) of the web citations linked to websites that contained the authors intended information but which were obviously updated. 143 (11%) web citations followed websites that were completely irrelevant to the

authors intended information. 85 (7%) web citations immediately returned the “Temporarily Unavailable” message and never became available during this study and thus remained in this category.

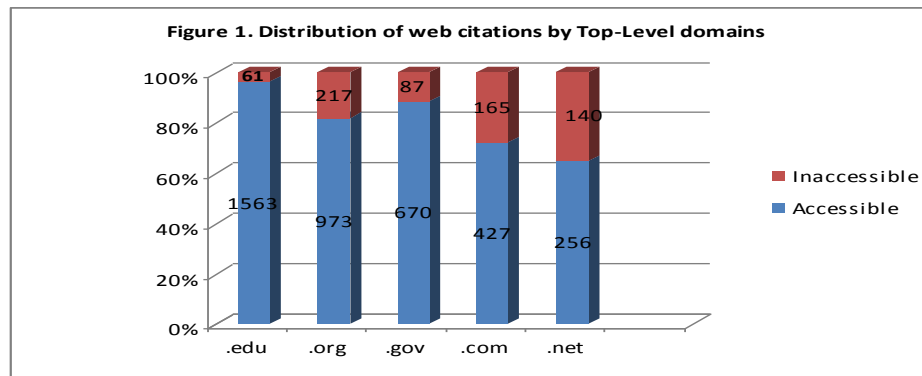
Table 3

*Frequency of Web Citations by Category in LIS Journals (2006-2010)*

<b>Internet Reference Category</b>	<b>2006 (n=475)</b>	<b>2007 (n=629)</b>	<b>2008 (n=758)</b>	<b>2009 (n=1137)</b>	<b>2010 (n=1254)</b>	<b>All Years (n=4253)</b>
Information readily accessible						
Active site	144(11.2)	378(8.9)	476(11.2)	721(17)	804(19)	2523(59.3)
Reference found with site search	113(2.7)	91(2.14)	47(2.6)	23(0.6)	75(1.8)	349(8.2)
Temporarily unavailable but now active	12(0.3)	0	75(1.8)	18(0.4)	22(0.5)	127(3)
Information not readily accessible						
Reference not found with site search	3(0.07)	7(0.1)	0	15(0.3)	6(0.1)	31(0.7)
No site search capability	8(0.1)	0	1(0.02)	5(0.1)	9(0.2)	23(0.5)
Temporarily unavailable	17(0.4)	6(0.1)	29(0.6)	13(0.3)	20(0.4)	85(2)
Updated version of same information	22(0.5)	40(0.1)	67(1.5)	56(1.3)	86(2)	271(6.4)
Information mismatch	23(0.5)	10(0.2)	18(0.4)	53(1.2)	39(0.9)	143(3.4)
No site found	133(3.1)	97(2.3)	45(1)	233(5.4)	193(4.5)	701(16.5)

As said by Maharana & Nayak (2006), a domain name can often tell the user if it is a government site, an academic site or a commercial site. In the present study, five different types of domain have been taken into consideration. They are .org, .edu, .com, .net and .gov; while those domains not falling into any of these categories fall into the “other” category.

The Top-level domain names of the web citations published in articles between 2006 and 2010 are shown in Figure 1. The .edu and the .org extensions comprise the highest number of URLs with a share of 1563 (37%) and 973 (23%), respectively. This indicates that the data sources of most of the web citations in the present study are websites of various educational institutions and professional institutions or societies, and the like. However, 670 cited URLs (16 percent) are of .gov type domains, followed by 427 (10 percent) from .com and 256 (6 percent) from .net and 364 (8 percent) from domains other than the above types.



The ratio for each top level domain in the 786 internet addresses that never became accessible was determined. Top-level domains associated with missing web located references are summarized in Figure 1. The top level domain having the greatest number of missing URLs was the organization domain (.org) – a finding that was also documented by Tan (2001). Disciplines that depend heavily on .org and .com sites will suffer the most from the effects of URL decay (Wagner, Gebremichael, Taylor & Soltys, 2009). Citations in .edu addresses were not only most common, they were also most reliable as 96% of .edu online citations worked. Next in terms of online permanence came citations in the .gov domain, which were accessible in 87% of the cases. It also supports a number of studies including those by Dimotrove & Bugeja (2009), Bar-Ilan & Pertiz (2009) and Tan (2001) that found URLs in the education domain to be the more persistent or stable than URLs in other domains.

### Discussion and Conclusion

There was a great disparity among disciplines in terms of citations to web resources. This disparity among different disciplines may be determined by the nature of the discipline, research methodology and differences in information use habits of individual researchers. The study revealed that there was a rising trend in using and citing web resources in the bibliographies of articles authored by LIS professionals in Iran. Thus, the rate of web citations has had an even and constant increase from 11% in 2006 to 30% in 2010 and this indicates the popularity and prevalence of web citations among Iranian researchers. This proves in LIS that attention to web resources in articles is significantly comparable with the reported findings from other researches (Isfandyari Moghaddam & Saberi, 2010; Wagner, Gebremichael, Taylor & Soltys, 2009; Aronsky, Madani, Carnevale, Duda & Feyder, 2007).

It is noteworthy that an about four year half-life was estimated for Iran's LIS publications, which is rather promising in comparison with other fields of study (Mardani, 2012; Isfandyari Moghaddam & Saberi, 2010). In a later four-year longitudinal study, Koehler (2002) found that the half-life of web pages equaled two years. He updated the

results from his earlier studies and concluded that webpage content becomes more stable over time, noting that the half-life of web resources in different disciplines, domains and fields differ (Koehler, 2004).

The information sources available in the .edu and .org domains are more research oriented and are obviously preferable to other websites and reflect the authors' perceived image towards these resources. Utilization of the .edu domain as the main source among online sources may reflect the authors' perceived image towards such sources because the provided information in these sources are suitable for citing purposes. Arguably, the .edu domain being associated with the knowledge-intensive university environments will tend to publish documents that are considered accurate and authoritative and exhibit a consistency that is generally associated with reputable educational entities (Sellitto, 2004). From 1999, the relative citation of sources located on .org servers increased significantly (Lawrence, 2001). The increased use of the .org domain may be due to the increased number of public and private associations now being online and making information more accessible to their constituencies. Indeed, authors may have discovered a perceived stability in resources located on the .org domain and their citation behavior reflects a confidence in using information appearing on this domain. In fact, the prevalence of these two mostly cited domains has been supported and shown in previous studies (Ducut, 2008; Mardani, 2012; Maharana & Nayak, 2006; Casserly & Bird, 2003).

The results of domains with content availability suggest that content at URLs with original domains of .edu and .gov is more likely to be permanent or accessible than is content located on other types of servers. Almost 96 percent of the content cited by URLs on educational servers was found at the URL cited. This result has been supported by Isfandyari Moghaddam & Saberi (2010), Dimitrova & Bugeja (2007) and Wren (2008). Therefore, these domains when compared to other ones prove more reliable and are less likely to disappear. Surprisingly, the results for the .org domain were different from findings from most of other studies (Isfandyari Moghaddam & Saberi, 2010; Dimitrova & Bugeja, 2007 and Maharana & Nayak, 2006). Nevertheless, there are other studies that have shown a great rate of inaccessible URLs in the .org domain (Sadat Mosavi, 2010). Our study provides further evidence that organization websites have become increasingly vulnerable to URL decay.

Along with the increasing use of web citations, the information accessibility problems have increased too. Some fields of study may be more prone to the effects of URL decay than others particularly if many of the scholarly materials utilized are available on the internet and rules permit the use of internet documents in scholarly materials. Web citation decay is a serious problem in LIS journals (Isfandyari Moghaddam & Saberi, 2010; Sadat mosavi, 2010). As a result, some researchers have suggested that web-based resources cannot be considered as a suitable research resource (Sellitto, 2005; Falagas, Karveli &

Tritsaroli, 2007). Clearly, it cannot be a fundamental solution to this problem and is rather a reaction out of helplessness. Since more than one fourth of the authors' intended information in web citations in this study was not accessible for a long time and about 18% of citations never became accessible, it is necessary for the Iranian authors to learn about using web citations as a source of information in their publications.

Because of the inconstant nature of URLs, neither publishers nor authors are able to guarantee the long-term accuracy or availability of online information referenced in LIS journals. Effective solutions will likely require a collaborative effort on the part of researchers, authors and journal editors.

Subsequently, several approaches of preservation of web content published in scholarly journals have been proposed, whether as policies and procedures (Johnson, 2004; Schilling, 2004), or computationally such as software tools (Kahle, 1997; Eysenbach, 2005; Reich & Rosenthal, 2004; Schafer, 2001) and unique tagging/tracking measures like digital object identifiers (DOIs) (Caplan, 1998).

In light of the limitations of URL preservation options, the importance of improving journal policies regarding URLs cannot be overstated. Zhang's (2011) study surveyed the editors of eight LIS journals and found that though they encouraged authors to cite web resources, they had only begun to work on policies relevant to this practice. Indeed, Zhang's (2011) review of the journal guidelines and instructions to the authors revealed an absence of clearly stated policies and/or guidelines regarding citing web resources.

It is also worth noting that in this study it was revealed that few of the Iranian LIS journals actually provide authors with instructions on citing web resources and this generally confirms the results of Zhang's (2001) editorial policy survey which revealed a "lack of clearly stated conventions on citing e-sources."

Editorial staff of LIS Journals in Iran should require authors to adhere to the citation policies, styles and formats established by their journals. Further, they should review their citation guidelines frequently and modify them, as needed, to ensure maximum access to the web content referenced by their authors. However, it is not clear to what extent the publishers and editorial staff of scholarly publications are concerned about the availability of cited electronic resources in the long run.

### References

- Aronsky, D.; Madani, S.; Carnevale, R. J.; Duda, S. & Feyder, M. T. (2007). The prevalence and inaccessibility of internet references in the biomedical literature at the time of publication. *Journal of the American Medical Informatics Association*, 14 (2), 232–234.
- Bar-Ilan, J. & Peritz, B. (2009). The lifespan of "informetrics" on the web: An eight year study (1998–2006). *Scientometrics*, 79 (1), 7–25.

- Caplan, P. (1998). DO I or Don't We? *PACS Review*. Retrieved on January 21, 2011 from <http://www.pr/v9/n1/capl9n1.html>.
- Casserly, M. & Bird, J. (2003). Web citation availability: Analysis and implications for scholarship. *College and Research Libraries*, 64 (4), 300–317.
- Dimitrova, D. & Bugeja, M. (2009). Exploring the use of online citations in an online-only journal: A case study of the journal of computer-mediated communication. *The Electronic of Communication*, 19 (1&2), 269-283.
- Dimitrova, D. & Bugeja, M. (2007). The half-life of internet references cited in communication journals. *New Media & Society*, 9 (5), 811–826.
- Ducut, E. (2008). An update on uniform resource locator (URL) decay in MEDLINE abstracts and measures for its mitigation. *BMC Medical Informatics and Decision Making*, 8 (23), DOI: 10.1186/1472-6947-8-23. Retrieved on February 6, 2010 from <http://www.biomedcentral.com/1472-6947/8/23/>.
- Eysenbach, G. (2005). Going, going, still there: Using the WebCite service to permanently archive cited web pages. *Journal of Medical Internet Resources*, 7(5), 60- 68.
- Falagas, M.; Karveli, E. & Tritsaroli, V. (2007). The risk of using the internet as reference resource: A comparative study. *International Journal of Medical Informatics*, 77 (4), 280-286.
- Harter, S. & Kim, H. (1996). Accessing electronic journals and other e-publications: An empirical study. *College & Research Libraries*, 57 (5), 440–456.
- Hester, S. (2004). Internet citations in oncology journals: A vanishing resource? *Journal of the National Cancer Institute*, 96 (12), 969–971.
- Isfandyari Moghaddam, A. & Saberi, M. (2010). Availability and half-life of web references cited in information research journal: A citation study. *International Journal of Information Science and Management*, 8 (2), 57-75.
- Johnson, K. (2004). Addressing internet reference loss. *Lancet*, 363 (9409), 660-661.
- Kahle, B. (1997). Preserving the Internet. *Scientific American*, 276 (3), 82-83.
- Koehler, W. (2002). Webpage change and persistence – A four-year longitudinal study. *Journal of the American Society for Information Science and Technology*, 53 (2), 162–171.
- Koehler, W. (2004). A longitudinal study of web pages continued: A report after six years. *Information Research*, 9 (2), 53- 71.
- 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.
- Lawrence, S. (2001). Persistence of web references in scientific research. *IEEE Computer*, 34 (2), 26–31.
- Maharana, B. & Nayak, K. (2006). Scholarly use of web resources in LIS research: A

- citation analysis. *Library Review*, 55 (9), 598-607.
- Mardani, A. (2012). An investigation of the web citations in Iran's chemistry articles in SCI. *Library Review*, 61(1), 18- 29.
- Markwell, J. & Brooks, D. (2002). Broken links: The ephemeral nature of educational www hyperlinks. *Journal of Science Education and Technology*, 11 (2), 105–8.
- Reich, V. & Rosenthal, D. (2004). Preserving today's scientific record for tomorrow. *BMJ Journal*, 328 (7431), 61-62.
- Roshandel, D.; Rezailashkajani, M.; Ansari, Sh. & Zali, M. (2005). Internet use by a referral gastroenterology clinic population and their medical information preferences. *International Journal of Medical Informatics*, 74(6), 447- 459.
- Sa'adat, A. (2008). Analysis of web citations in library and information science journals. *Fasnam-e Ketab*, 18 (1), 41-50.
- Sadat Mosavi, A. (2010). Assessing accessibility in four library and information sciences journals: A comparative study. *Information Sciences & Technology*, 25 (4), 617-633.
- Sellitto, S. (2005). The impact of impermanent web-located citations: A study of 123 scholarly conference publications. *Journal of the American Society for Information Science and Technology*, 56 (7), 695–703.
- Sellitto, S. (2004). A study of missing web-cites in scholarly articles: Towards an evaluation framework. *Journal of Information Science*, 30 (6), 484- 495.
- Schafer, K. (2001). The PURL project. *Journal of Library Administration*, 34 (1-2), 123-149.
- Schilling, L. (2004). Bioinformatics leads charge by publishing more internet addresses in abstracts than any other journal. *Bioinformatics*, 20 (7), 2903-2919.
- Spinellis, D. (2003). The decay and failures of web references. *Communications of the ACM*, 46 (1), 71-77.
- Tan, A. (2001). FOCI: A personalized web intelligence system. *Proceedings of the IJCAI Workshop on Intelligence Techniques for Web Personalization*, 14-19.
- Thorp, A. & Brown, L. (2007). Accessibility of internet references in annals of emergency medicine: Is it time to require archiving? *Annals of Emergency Medicine*, 50 (2), 188-193.
- Tyler, D. & McNeil, B. (2003). Librarians and link rot: A comparative analysis with some methodological considerations. *Portal: Libraries and the Academy*, 3 (4), 615- 632.
- Wagner, C.; Gebremichael, M.; Taylor, M. & Soltys, M. (2009). Disappearing act: Decay of uniform resource locators in health care management journals. *Journal of the Medical Library Association*, 97 (2), 122-130.
- Wren, J. (2008). URL decay in MEDLINE: A 4-year follow-up study. *Bioinformatics*, 24 (11), 1381-13855.
- Zhang, Li. (2011). Use of web resources in the journal literature 2001 and 2007: A cross-

disciplinary study. *College & Research Libraries*, 72 (2), 167-179.

Zhang, Y. (2001). Scholarly use of internet-based electronic resources. *Journal of the American Society for Information Science and Technology*, 52 (8), 628-54.