Annals of Library and Information Studies
Vol. 66, March 2019, pp. 24-32

Recovery of missing URLs cited in *Annals of Library and Information Studies*: a study of Time Travel

D. Vinay Kumar[a] and M. Sushmitha[b]

[a]Lecturer, Department of Library and Information Studies, Kuvempu University, Jnanasahyadri, Shankaraghatta, Shivamogga–577451, Email: vinay.86.kumar@gmail.com

[b]Department of Library and Information Studies, Kuvempu University, Jnanasahyadri, Shankaraghatta, Shivamogga–577451

Received: 19 July 2018; revised and accepted: 26 March 2019

This paper intends to know the rate of active and missing URLs cited in the article published in *Annals of Library and Information Studies* during 2006-2015. A total of 6713 references appended to 342 research articles published during ten years were studied. Of the 6713 references, 1105 have URLs. It is found that 56.56% of URLs are active and remaining 43.44% of articles are missing. 'HTTP-404' is the common error code associated with missing URLs. The paper also discusses the usefulness of Time Travel in the recovery of missing URLs. The results show that Time Travel was able to recover more than 50% of missing URLs through different web archives.

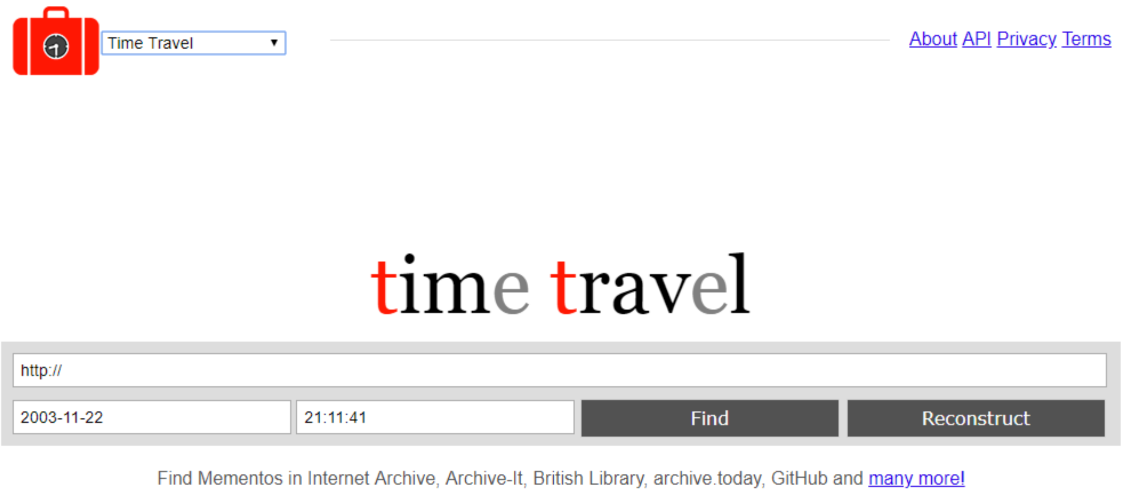
Keywords: URL; References; Journal

Introduction

The Internet, a very complex and revolutionary invention that plays a key role in rapid access to information. The growing e-content on the web has made it the quickest and appropriate medium for global information exchange[1]. The adoption of web-based technology for publishing information has quickened and increased the availability of electronic information over the web and on the other hand, it has influenced the large use of web resources in scholarly literature[2]. With the advancement of the Internet, researchers have the privilege to access full-text information available on the web. Therefore, the web is popular among researchers and it is as an important means of locating and sharing scientific information. The present-day researchers fraternity prefer to access information sources on the web[2]. They have been increasingly citing Universal Resource Locators (URLs) in their scholarly publications. Library and information science is no exception to this phenomenon[3].

One major disadvantage with web resources is that they subject to modifications and they can be overwritten[4]. This has lead researchers to investigate the permanency of URLs of web resources cited in scholarly works. Every researcher expects that the URL citations should remain stable and accessible permanently. Therefore, a mechanism is required to monitor the web resources, which also preserve them for future use[5]. Correspondingly, various web-archiving initiatives have come up to preserve the web-based knowledge permanently. The web archives store web documents permanently. Thus, they ensure long-term access to the URLs that are not accessible through the web browsers. These web archives are also being used to recover the URLs that are no longer available on the web. Time Travel is such a tool that can be used to recover the missing URLs.

Time Travel (http://timetravel.mementoweb.org/) is a free tool that recovers the missing URLs. When a missing URL is submitted to the search box of the Time Travel and the find button is clicked, it locates the missing URL (Fig.1) and shows the name of web archive from where the submitted missing URL is recovered (Fig. 2). The Time Travel service recovers the missing URLs that are archived in any one of these web archives viz., Archive Today, Archive-It, Bibliotheca Alexandrina Web Archive, DBpedia archive, DBpedia Triple Pattern Fragments archive, Canadian Government Web Archive, Croatian Web

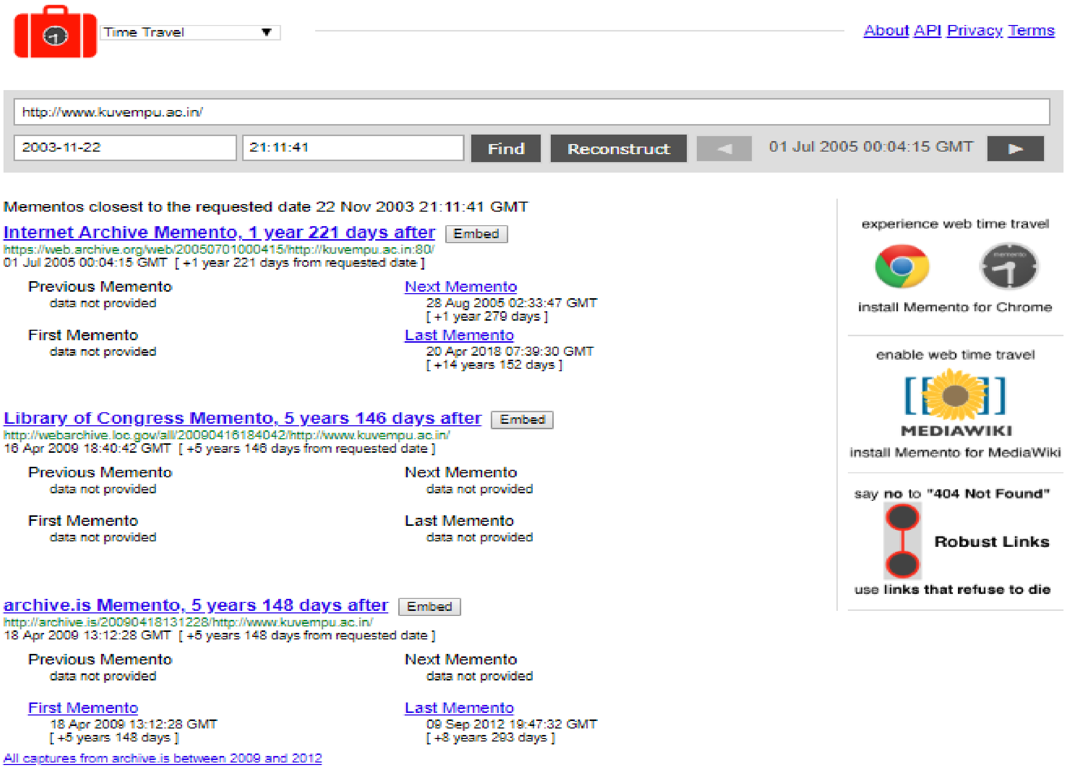


http://

2003-11-22 21:11:41

Find Mementos in Internet Archive, Archive-It, British Library, archive.today, GitHub and many more!

Fig. 1—Homepage of Time Travel


Time Travel
About API Privacy Terms
Find
Reconstruct
Embed
experience web time travel
install Memento for Chrome
enable web time travel
MEDIAWIKI
install Memento for MediaWiki
say no to "404 Not Found"
Robust Links
use links that refuse to die

http://www.kuvempu.ac.in/

2003-11-22 21:11:41 01 Jul 2005 00:04:15 GMT

Mementos closest to the requested date 22 Nov 2003 21:11:41 GMT

Internet Archive Memento, 1 year 221 days after
https://web.archive.org/web/20050701000415/http://kuvempu.ac.in:80/
01 Jul 2005 00:04:15 GMT [+1 year 221 days from requested date]

Previous Memento
data not provided

Next Memento
28 Aug 2005 02:33:47 GMT
[+1 year 279 days]

First Memento
data not provided

Last Memento
20 Apr 2018 07:39:30 GMT
[+14 years 152 days]

Library of Congress Memento, 5 years 146 days after
http://webarchive.loc.gov/all/20090416184042/http://www.kuvempu.ac.in/
16 Apr 2009 18:40:42 GMT [+5 years 146 days from requested date]

Previous Memento
data not provided

Next Memento
data not provided

First Memento
data not provided

Last Memento
data not provided

archive.is Memento, 5 years 148 days after
http://archive.is/20090418131228/http://www.kuvempu.ac.in/
18 Apr 2009 13:12:28 GMT [+5 years 148 days from requested date]

Previous Memento
data not provided

Next Memento
data not provided

First Memento
18 Apr 2009 13:12:28 GMT
[+5 years 148 days]

Last Memento
09 Sep 2012 19:47:32 GMT
[+8 years 293 days]

All captures from archive.is between 2009 and 2012

Fig. 2—Recovery of URLs through Time Travel

Archive, Estonian Web Archive, Icelandic web archive, Internet Archive, Library of Congress Web Archive, NARA Web Archive, National Library of Ireland Web Archive, perma.cc, Portugese Web Archive, PRONI Web Archive, Slovenian Web Archive, Stanford Web Archive, UK Government Web Archive, UK Parliament's Web Archive, UK Web Archive, Web Archive Singapore, WebCite.

Review of literature

Many studies have focused on the trend, accessibility issues of the URL citations cited in LIS scholarly literature and found the problem of missing URLs, and have examined the recovery of missing URLs using web archives and Google search engine. A study by Moghaddam and Saberi (2010) examined the availability and half-life of URLs cited in articles published by *Information Research* journal. To do this, at first, they have taken all issues of *Information Research* from April 1995 to March 2008 and counted all the citations. Thereafter the authors checked the accessibility of the individual cited URLs. When they could not access a URL in any article, they tried to visit the referred website. If this attempt seemed to be inadequate, search engine "Google" was employed to access the missing reference(s). The 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, in 1761 cited URLs, 73% were accessible, and 27% were inaccessible. It is notable that methods adopted to recover the missing URLs such as using Google, search for URLs have increased the accessible URLs increased from 73% to 86%[6].

Mardani (2011) studied 46762 citations (20%) of the total 187823 available citations in the articles included web citations. The proportion percentage of web citations increased from 9% in 2006 to 39% in 2009. The average number of web citations for every article is 4.52. The most widely cited top-level domains in URLs include the .org and .edu with 31% and 23% respectively, and when compared to other domains they reveal a greater tendency for stability. The highest percentage of inactive URLs was found to be associated with the .gov top-level domain. Ultimately, 40954 web citations were rendered accessible, of which 79% allowed easy and long-term access to the author's information intended in URLs. The decay rate for citations reveals an annual 5.2% increase[7].

Sampath Kumar and Manoj Kumar (2012) investigated the decay of online citations cited in four open access journals published between 2000 and 2009. The study found that a total of 1158 online citations cited in 1086 research articles published in two science and social science journals spanning a period of 10 years (2000-2009) were extracted. The study found that 24.58% (267 out of 1086) of articles had online citations and 30.56% (26% in Science and 52.73% in Social Science) of online citations were not accessible and remaining 69.44% of online citations were still accessible. The 'HTTP 404 error message-page not found' was the common message encountered and represented 67.79% of all HTTP message. Domains associated with .ac and .net had higher successful access rates while .org and .com/.co had the lowest successful access rates[8].

Mardani and Sangari (2012) studied the current situation and characteristics of web citations accessibility. The 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, being 37% and 23% respectively. The results show that only 3467 web citations remained accessible in 2011, of which 71% allowed easy and long-term access to the authors‟ information intended in URLs. Longtime 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[9].

Prithviraj and Sampath Kumar (2014) studied and analyzed the accessibility, and corrosion of URLs cited in the articles of Indian LIS conference proceedings published from 2001 to 2010. A total of 5,698 URLs cited in the 1,700 articles were examined. The percentage of URLs increased from 39.10 percent in 2001 to 73.47 percent in 2009. The study found that 50.09 percent of URLs were not accessible at the time of testing and the remaining 49.91 percent of URLs were accessible. The HTTP 404 error message – ‘‘file not found’’ was the common error message encountered and represented 53.29 percent of all HTTP messages[10].

Gul et al. (2014) have made an attempt to explore and analyze the growth and decay rate of URL citations cited in one of the eminent information web magazine *ARIADNE*. All the web citations cited in the featured article section of *ARIADNE* web magazine spanning a period of three years (2010-2012) were identified and downloaded. All the URLs provided in the web citations were checked by W3C link checker (http://validator.w3.org/checklink) to test whether they were accessible or not. They found that Web resources are used more than the printed ones. Though the number of references in web-based format is more in use, a significant number of web references lose their identity with the passage of time. The results indicate that early-published papers have collectively a larger number of missing web citation compared to the recent ones. It was found that the majority of errors were due to the missing content (HTTP 404-file not found) representing 52.68% of all HTTP error codes followed by "HTTP 500" (24.73%). The ".com/.co" domain was also found to be the most stable and persistent domain with 95 % accessibility. The greatest number of web resources cited in the articles were found to be of "HTML" and "PPT" files were found to be most stable with 100% accessibility. Finally, the study confirmed that continued availability of web resources is not guaranteed because of the dying phenomenon of web-based references, but there can be certain solutions (WebCite, LOCKSS etc.) that can prevent the decay or disappearance of web citations[11].

Sampath Kumar et al. (2015) studied the URLs to know the rate of loss of online citations used as references in scholarly journals. It also indented to recover the vanished online citations using Wayback Machine. The study selected three journals published by Emerald publication. All 389 articles published in these three scholarly journals were selected. A total of 15,211 citations were extracted of which 13,281 were print citations and only 1,930 were online citations. The study found that 30.98 percent of them were not accessible. The HTTP 404 error message –"page not found" was the common message encountered and represented 62.98 percent of all HTTP error message. It was found that the Wayback Machine had archived only 48.33 percent of the vanished webpages[12].

Tajeddini et al. (2018) compared currency, disappearance and half-life of URLs cited in Iranian research articles indexed in ISI in three disciplines during 2009-2011. The results indicated that the three disciplines viz., information science, psychology, and management had 13.7%, 44.8%, and 14.23% of URL citations respectively. At the initial check, they found that 25% of overall URLs were inaccessible. They used the Internet Archive and Google to recover the inaccessible URLs which reduced the inaccessible URLs from 25% to 3%[13].

Kumar and Kumar (2017) studied the accessibility and permanency of URLs cited in the articles published in *DESIDOC Journal of Library and Information Technology* during 2006-2015. A total of 7353 citations were cited in 491 articles of which 2133 were URLs. The study found 823 (38.58%) URLs were not accessible and remaining 1310 (61.42%) were accessible. HTTP 404 was the highest error message associated with missing URLs. The study used the Internet Archive to recover the missing URLs and able to recover 58.81% of them[14].

Sife (2017) studied the availability and persistence of URLs cited in health science journals published during 2001-2015. The study reported that of the 17,609 citations only 574 (3.3%) were URL citations of which 253 (44.1%) were not accessible. HTTP-404 was again the common message encountered with the missing URLs. He used Wayback Machine to retrieve the missing URLs and was able to recover only 6.3% of missing URLs[15].

While there are many studies on the permanency of URLs, the use of multiple web archives to recover the vanished URLs have not been extensively attempted by researchers. Majority of studies used the *Internet Archive* as the recovery tool. However, the use of *Time Travel* as a recovery tool would retrieve the data from many web archives including the *Internet Archive.* Hence, this study focused on the use of *Time Travel* in the recovery of missing URLs cited in the articles published in *Annals of Library and Information Studies*.

Objectives of the study

- To assess extent of use of URL citations in and Indian LIS journal;
- To examine the diversity of top-level domains and file formats associated with total, missing, and recovered missing URLs;
- To identify HTTP error messages associated with missing URLs; and
- To know the extent of recovery of missing URLs through Time Travel

Methodology

The study explored the frequency and accessibility of URL citations cited in a total of 342 research articles published in *Annals of Library and Information Studies* during 2006-2015. The period for the analysis was the publication years 2006-2015. The selection of this period intends to include the timebound URLs cited in scholarly literature. All research articles published in ALIS journal during the 10 years were downloaded. A total of 342 research article consist of 6713 references were considered for the study. Of the 6713 unique references, 1105 were URL citations.

Testing of URLs

Once the extraction of URLs from the reference list was completed, the URLs were tested for their accessibility using W3C Link Checker (http://validator.w3.org/checklink). This tool tests a submitted URL for broken or non-valid hypertext links. A useful feature of this application was that if a broken link was found, it brings out the exact HTTP error message. Upon completion of the testing of URLs using W3C link checker, the inaccessible URLs have been considered as 'missing URLs' and the exact HTTP error message was recorded before those missing URL. The accessible URLs were considered as 'active URLs'.

Recovery of missing URLs

The URL citations which gave HTTP error messages were entered one by one into the search box of the Time Travel and searched. Time Travel recovered the missing URL citations and shows the history of archived the missing web sources.

Analysis

Table 1 shows that the 342 research articles published in *Annals of Library and Information Studies* during the years 2006-2015 consisted of 6,713 references of which 1,105 had URLs in the references (16.46%). The percentage of URL references varied from a low of 5.63 percentages in the year 2009 to a high of 29.68% in the year 2012.

This shows that the volume of URL citations in the journal articles was not consistent. Although previous studies (Dimitrova and Bugeja, 2007[16]; Sampath and Manoj, 2012[8]; Sife and Bernard, 2013[17]) have documented that internet sources have emerged as the most popular sources among the research community, the present study did not follow this trend.

Distribution of active and missing URLs

The URLs were checked using the W3C link checker to know whether the web source could be accessible at the given URL that was used in the references.

Table 2 shows that out the 1,105 URLs, 56.56% of URLs are active and 43.44% of URLs are missing. The percentage of missing URLs varied from 30.17% in the year 2015 to the highest of 78.95% in the year 2006.

HTTP Error-wise distribution of missing URLs

There are many reasons for failure of URLs and the type of errors may be HTTP 301, HTTP 400, HTTP 403, HTTP 404, HTTP 410, HTTP 500 and HTTP 503.

Table 3 shows that the HTTP 404 error message—"page not found" was the overwhelming message encountered and represented 75.63 percent of all HTTP error message and it is followed by HTTP 500 (14.38%) and HTTP 301 (2.50%).

Table 1—Distribution of articles, citations and URL citations by year

Year	Total articles	Total citations	Percentage	Total URLs	Percentage
2006	26	385	5.74	57	14.81
2007	28	384	5.72	29	7.55
2008	35	591	8.80	121	20.47
2009	34	675	10.06	38	5.63
2010	43	1039	15.48	201	19.35
2011	36	819	12.20	134	16.36
2012	27	465	6.93	138	29.68
2013	37	837	12.47	110	13.14
2014	40	907	13.51	161	17.75
2015	36	611	9.10	116	18.99
Total	342	6713	100.00	1105	16.46

Table 2—Distribution of active and missing URL citations

Year	Total URLs	Active URLs	Percentage	Missing URLs	Percentage
2006	57	12	21.05	45	78.95
2007	29	17	58.62	12	41.38
2008	121	62	51.24	59	48.76
2009	38	15	39.47	23	60.53
2010	201	106	52.74	95	47.26
2011	134	80	59.70	54	40.30
2012	138	91	65.94	47	34.06
2013	110	76	69.09	34	30.91
2014	161	85	52.80	76	47.20
2015	116	81	69.83	35	30.17
Total	1105	625	56.56	480	43.44

Table 3—HTTP error of missing URLs

HTTP Error	Missing	Percentage
HTTP – 301	12	2.50
HTTP – 400	2	0.42
HTTP – 403	28	5.83
HTTP – 404	363	75.63
HTTP – 410	2	0.42
HTTP – 500	69	14.38
HTTP – 503	1	0.21
Total	480	0.63

Recovery of missing URLs through Time Travel

The study intended to recover the vanished URLs using the Time Travel tool. All 480 missing URLs were entered in the search box of 'Time Travel' by copying the exact URL and the searched . The results are presented in Table 4.

Table 4 shows that Internet Archive recovered the highest percentage of missing URLs (56.46%) followed by Bibliotheca Alexandria Web Archive (15%). Other web archives have recovered less than 5% of missing URLs.

Table 5 shows the distribution of total and missing URLs by their domain. Of the 1105 URLs, the highest percentage of URLs belonged to the .org domain (31.04%) followed by .co/.com domain (24.43%) (and .edu domain (13.30%). The table also shows that the domain having the greatest number of missing URLs was .ac with 67.05% followed by .int (54.55%) and .edu (53.06%).

A notable finding is that none of the DOIs and URLs belonged to the domain .info have missed. This clearly indicates the permanency of DOIs as citations.

The data presented in Table 6 shows that the highest percentage of URLs lead to HTML files (70.68%). PDF files were next that accounted for 20.81%. It is evident from the data that more than 90% of the URL citations are HTML and PDF files.

The data in Table 6 also indicates the missing URLs by file formats. File formats having the highest percentage of missing URLs is .doc (83.33%) followed by .ppt (66.67%). The data also reveals that .htm/.html and .php file formats are more stable than other file formats accounting for 39.18% and 41.03% of the loss.

Table 7 shows the top-level distribution of recovered URL citations by their domains. It is evident from the table that Time Travel recovered the missing URLs that belonged to different domains. Of the 135 missing URLs with the .org domain, 104 were recovered by Internet Archive, which is followed by .co/.com (55 out of 88), .edu (43 out of 78) (and country code domain (19 out of 61). Another web archive, Bibliotheca Alexandria could able to recover 39 missing URLs with .org domain followed by 7 missing URLs with the .gov domain. Other web archives recovered very few numbers of missing URLs.

The distribution of recovered URL citations by file format is given in Table 8. It is clear that missing URLs with HTML file format have been highly recovered from all archives followed by the missing URLs with PDF file format. Among the various web archives, Internet Archive recovered the highest number of missing URLs with different file formats.

Conclusion

Citations to web sources have been increasing in scholarly literature. Meanwhile, the problem of missing URLs still persist and it can be addressed

Table 4—Distribution of recovered missing URLs through Time Travel

		Recovered from							
Year	Missing URLs	Internet Archive	Bibliotheca Alexandrina Web Archive	Libraryof Congress	archive.is	Archive-It	Portuguese Archive	UK Web Archive	perma.cc
2006	45	36	11	6	2	3	3	0	0
2007	12	7	2	1	2	2	5	1	0
2008	59	47	8	3	3	2	3	0	1
2009	23	12	5	3	1	2	1	0	0
2010	95	63	13	4	1	1	0	0	0
2011	54	29	8	2	2	1	0	0	0
2012	47	26	3	1	2	1	1	0	0
2013	34	16	6	1	5	1	0	0	0
2014	76	26	12	2	1	2	1	0	0
2015	35	9	4	1	1	0	0	0	0
Total	480	271	72	24	20	15	14	1	1
Percentage		56.46	15.00	5.00	4.17	3.13	2.92	0.21	0.21

Table 5—Domains associated with missing URLs

Domains	Total	% to total URLs	Missing	%
.org	343	31.04	135	39.36
.co/com	270	24.43	88	32.59
.edu	147	13.30	78	53.06
Geo domains	92	8.33	61	66.30
.ac	88	7.96	59	67.05
.gov	57	5.16	28	49.12
.net	32	2.90	13	40.63
.nic	19	1.72	5	26.32
.res	16	1.45	3	18.75
.ernet	13	1.18	4	30.77
.int	11	1.00	6	54.55
Only DOI	10	0.90	0	0.00
.info	7	0.63	0	0.00
Total	1105	100.00	480	43.44

Table 6—File formats associated with missing URLs

File Formats	Total	% to total URLs	Missing	Percentage
HTML/HTM	781	70.68	306	39.18
PDF	230	20.81	125	54.35
PHP	39	3.53	16	41.03
ASP	28	2.53	16	57.14
CFM	8	0.72	5	62.50
CGI	7	0.63	4	57.14
DOC	6	0.54	5	83.33
PPT	3	0.27	2	66.67
JSP	2	0.18	1	50.00
NSF	1	0.09	-	0.00
Total	1105	100.00	480	43.44

Table 7—Domains associated with recovered URLs

Domain	Missing URLs	Recovery of missing URLs through							
		Internet Archive	Bibliotheca Alexandrina Web Archive	Library of Congress	archive.is	Archive-It	Portuguese Archive	UK Web Archive	perma.cc
.org	135	104	39	7	4	9	4	0	0
.co/.com	88	55	6	1	2	2	5	1	0
.edu	78	43	5	3	3	1	3	0	0
.ac	59	18	3	3	1	0	1	0	0
.nic	5	5	3	4	0	0	0	0	0
.net	13	2	0	0	0	0	0	0	0
.ernet	4	3	0	0	0	0	0	0	0
Country code	61	19	6	0	5	0	0	0	0
.int	6	5	2	2	1	2	1	0	0
.res	3	2	1	0	1	0	0	0	0
.gov	28	15	7	4	3	1	0	0	1
Total	480	271	72	24	20	15	14	1	1

Table 8—File formats associated with recovered URLs

File Format	Total missing URLs	Recovered of missing URLs through							
		Internet Archive	Bibliotheca Alexandrina Web Archive	Library of Congress	archive.is	Archive-It	Portuguese Archive	UK Web Archive	perma.cc
.htm/.html	306	206	69	20	19	8	11	1	1
.pdf	125	55	3	4	0	6	3	0	0
.doc	5	1	0	0	0	0	0	0	0
.ppt	2	1	0	0	0	0	0	0	0
.cgi	4	0	0	0	0	0	0	0	0
.php	16	6	0	0	1	1	0	0	0
.asp	16	2	0	0	0	0	0	0	0
.jsp	1	0	0	0	0	0	0	0	0
.cfm	5	0	0	0	0	0	0	0	0
Total	480	271	72	24	20	15	14	1	1

with the use of recovery tools such as Web Archives in general and Time Travel. Though the web archives crawl the web and archives the web resources, they have limitations such as copyright issues, denying the permission to capture the content of the website etc. Hence, there is a need to strengthen the web archives by means of upholding the issues of fair use and permanent preservation of web content. At the same time, the webmasters and the creators of web resources need to permit the web archives to crawl their sites so that a scholarly content will be saved permanently.

References

1. Riahinia N, Zandian, F and Azimi Ali, Web citation persistence over time: a retrospective study, *The Electronic Library,* 29(5) (2010) 609-620.

2. Spinellis D, The decay and failures of web references, *Communication of the ACM*, 46(1) (2003)71–77.

3. Maharana B, Nayak Kand Sahu NK, Scholarly use of web resources in LIS research: a citation analysis, *Library Review,* 55(9) (2006) 598-607.

4. Casserly MF and Bird JE, Web citation availability: Analysis and implications for scholarship, *College and Research Libraries,* 64(7) (2003) 300–317.

5. `Wren JD, 404 not found: the stability and persistence of URLs published in MEDLINE, *Bioinformatics*, 20(5) (2004) 668-672.
6. Moghaddam IA and Saberi MK, Availability and half-life of web references cited in information research journal: A citation study, *International Journal of Information Science and Management,* 8(2) (2010) 57–75.
7. Mardani A, An investigation of the web citations in Iran's chemistry articles in SCI, *Library Review,* 61(1) (2011) 18-29.
8. Sampath Kumar BT and Manoj Kumar KS, Persistence and half-life of URL citations cited in LIS open access journals, *ASLIB Proceedings*, 64(4) (2012) 405-422.
9. Mardani A and Sangari M, The Availability and Persistence of Web Citations in Iranian LIS Journals (2006-2010), *Library Philosophy and Practice*, 2012, Retrieved from http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1889andcontext=libphilprac
10. Prithvi Raj KR and Sampath Kumar BT, URLs as references in Indian LIS conference papers: an Analysis, *Annals of Library and Information Studies,*60(4) (2014)284-295.
11. Gul S, Mahajan I and Ali A, The growth and decay of URLs citation: A case of an online Library and Information Science Journal, *Malaysian Journal of Library and Information Science,* 19(3) (2014) 27-39.
12. Sampath Kumar BT, Vinay Kumar D and Prithviraj KR, Wayback Machine: reincarnation to vanished online citations, *Program*, 49(2) (2015) 129-141.
13. Tajedini O, Ghazizade A and Tajedini A, Investigation of the currency, disappearance and half-life of URLs of web resources cited by Iranian Researchers: a comparative study, *International Journal of Information Science and Management*, 16(1) (2018) 27-47.
14. Vinay Kumar D and Sampath Kumar BT, Finding the unfound: Recovery of missing URLs through Internet Archive, *Annals of Library and Information Studies,* 64(3) (2017) 165-171.
15. SifeAS and Lwoga ET, Retrieving vanished Web references in health science journals in East Africa, *Informationand Learning Sc*ience, 118(7/8) (2017) 385-392.
16. Dimitrova DV and Bugeja M, The half-life of Internet references cited in communication journals, *New Media and Society*, 9(9) (2007) 811–826.
17. Sife AS and Bernard R, Persistence and decay of web citations used in theses and dissertations available at the Sokoine National Agricultural Library, Tanzania*, International Journal of Education and Development using Information and Communication Technology,* 9(2) (2013) 85.