Persistence and decay of web citations used in theses and dissertations available at the Sokoine National Agricultural Library, Tanzania

Alfred S. Sife and Ronald Bernard,
Sokoine University of Agriculture, Tanzania

ABSTRACT

A study was conducted to examine the persistence and decay of web citations in theses and dissertations available at the Sokoine National Agricultural Library. Specifically, the study assessed the accessibility status of cited URLs, identified error messages and top level domains of inaccessible URLs, and calculated the half-life of web citations. Eighty-three theses and dissertations that were dated between 2007 and 2011 were stratified according to their years of publication and randomly selected for the study. These gave a total of 15,468 citations of which 1,487 (9.6%) were web citations. The findings show that a total of 862 (58%) web citations were inaccessible. The *404 File Not Found* error message was the most (92.7%) encountered and the *.com* domain had the greatest number (28.2%) of missing URLs. The average half-life for the URLs cited in theses and dissertations was 2.5 years. The study findings therefore indicate that many web resources cited in theses and dissertations available at SNAL had disappeared from their original locations. Collaborative efforts are thus required from various stakeholders in order to reduce the problem of URL decay.

Keywords: web citations, URL persistence, URL decay, theses, dissertations, Tanzania

INTRODUCTION

Scholarly writing requires that authors make reference to the previously published works by mentioning the authors inside their works (in-text citation) and giving bibliographic details in the lists of references. That is to say newer scholarly works are supposed to cite the older ones; which is an important characteristic that distinguishes academic research writing from other kinds of writings (Krause, 2007). Citations enable authors to retrieve publications again, support and substantiate their arguments and claims, build upon existing works, credit other authors' ideas, acknowledge intellectual indebtedness, provide the context in which research is performed, and compare different approaches and methodologies. Citations also signal authors' awareness of ethical publishing principles, enables readers to explore further any topic of interest, and determines the popularity and impact of specific publications and authors (Ticehurst and Veal, 2000; Webster and Watson, 2002). Furthermore, citations show that the advancement of knowledge is incremental with new investigations relying on previous works to produce new knowledge. This kind of relationship has been termed as *"authors standing on the shoulders of others in furthering their own works"* (Sellitto, 2004).

Following the phenomenal growth in the web-based information which has altered the *modus operandi* of scholars by providing new avenues for scholarly communication, many citation conventions, styles and guidelines require authors to cite Uniform Resource Locators (URLs) as part of the bibliographic details in the lists of references. A URL is an address of the location of an electronic document on the Web consisting of four parts - protocol, domain, directory and file. A Web *protocol*, which is also called the Hypertext Transfer Protocol (HTTP), is a set of communication rules for exchanging information and enables browsers to connect with web servers. A *domain* is the way to identify and locate computers connected to the Internet. The last part of the domain name, called the top level domain, can tell the type of organization such as

.com for commercial organizations, *.edu* or *.ac.* for educational organizations, *.gov* for government sites and *.org* for non-profit organizations (Tajeddini *et al.,* 2011). *Directory* is the name on the server for the folder from which the browser needs to pull the *file* that contains information. Increasing volumes of web-based information coupled with demands to include URLs in the references have led to increased web citations in scholarly publications (Germaine, 2000; Rumsey, 2002; Spinellis, 2003; Goh and Ng, 2007).

Technically, a URL must be resolved to a valid Internet Protocol (IP) address; otherwise an HTTP error message occurs. There are numerous errors that can be encountered (Table 1) but the most common one (*404 File Not Found)* occurs when the server has not found anything matching the requested URL. Other errors occur when the request cannot be understood by the server; user authentication required; server refuses to fulfil a request; server has not found anything matching the requested URL; and when the server cannot send an acceptable response according to the accept header. Errors also occur when a server encounters unexpected conditions which prevent it from fulfilling a request; server is unable to handle a request; and when a server does not receive a timely response from an upstream server specified by the URLs*.* An error message *901 Host lookup Failure* is encountered when the host name cannot map an IP address (Powell, 2003; Spinellis, 2003).

Table 1: *Types of errors that may be encountered*

Type of errors	Description
400 Bad request	Request cannot be understood by the server
401 Unauthorized	Request requires user authentication
403 Forbidden	Server refuses to fulfil a given request
404 Not Found	Server has not found anything matching the requested URL
406 Not Acceptable	The resource identified by the request is only capable of generating response entities which have content characteristics not acceptable according to the accept headers sent in the request.
410 Gone	The requested resource is no longer available at the server and no forwarding address is known.
500 Internal Server Error	Server encounters unexpected condition which prevents it from fulfilling the request
503 Service Unavailable	Server is unable to handle the request due to temporary overloading or maintenance
504 Gateway Time-out	Server does not receive timely response from the upstream server specified by the URL or some other auxiliary server needed to complete the request.
901 Host lookup Failure	The host name cannot map an IP address

Citing URLs in the lists of references is an academic requirement which stems from the assumption that a particular information resource will continue to be located at the cited URL. However, continued availability of web resources is often not guaranteed as they may disappear intermittently or permanently from their locations. Web resources disappear when original documents have been removed and their URLs changed; content has been altered; or equipment such as servers is down. Changes made to websites such as reconstruction, merging, redirecting and expansion can mean inconsistency in URLs. Another common reason for URLs failure is that they might be transcribed or typed incorrectly. The phenomenon associated with the disappearance of cited URLs has been coined as "broken links", "ephemeral nature of web hyperlinks", "link rot", "missing web-cites", "going 404", "web decay" or "URL decay" (Markwell

and Brooks, 2002; 2003; Sellitto, 2004; Wren, 2008). Failure to locate online references not only undermines the foundations of academic writing but also raises questions on the practical implications of citing URLs.

While numerous studies (Germaine, 2000; Lawrence *et al.,* 2001; Rumsey, 2002; Spinellis, 2003) have focused on the inaccessibility of cited URLs in scholarly journal articles, there is scarce research on the permanency of web references in students' works particularly postgraduate theses and dissertations. This study therefore examined the persistence and decay of URLs cited in theses and dissertations dated between 2007 and 2011 that were available at the Sokoine National Agricultural Library. This kind of study was considered important because theses and dissertations are similar to other scholarly publications in that they are based on original research and students are often bound to use scholarly information. Specifically, the study assessed the accessibility status of cited URLs, identified error messages associated with inaccessible URLs, identified the top domain levels of decayed URLs, and calculated the half-life (time required for half of all web citations in a publication to disintegrate) of the web citations referred in theses and dissertations.

LITERATURE REVIEW

Various studies (Harter and Kim, 1996; Koehler 1999; Davis and Cohen 2001; Koehler 2002; Casserly and Bird, 2003; Tyler and McNeil, 2003; Wren *et al.*, 2006; Dimitrova and Bugeja, 2007) have reported on the problem of inaccessible URLs. Harter and Kim (1996) examined 47 unique URLs from e-journals published during 1993-1995 and reported that 31% of the URLs were unavailable. Tracking 350 URLs from 1996 to 1998, Koehler (1999) found that 17.7% of websites and 31.8% of web pages failed to respond when queried after 12 months. In another study, Koehler (2002) examined the attrition and modification of websites/pages. This study confirmed many of the previous findings and indicated that the average web page half-life was approximately two years. Casserly and Bird (2003) examined 500 internet citations randomly chosen from scholarly articles published in library and information science (LIS) journals. They reported that 56.4% of those URLs were accessible, while the rest disappeared from the original web addresses and that *"File Not Found"* was the most frequent error message.

Investigating URL decay in dermatology journal articles published between 1999 and 2004, Wren *et al.* (2006) found that 81.7% of 1113 URLs were available but decreased with time since publication from 89.1% of 2004 URLs to 65.4% of 1999 URLs. Dimitrova and Bugeja (2007) studied cited URLs in journalism and communication field and reported that 39% of URLs were unavailable and the .org was the most available domain with 70% active links. Goh and Ng's (2007) study on accessibility and decay of URLs of three LIS journals during 1997-2003 revealed that the decayed rate of URLs was 31%. They also reported that 56% of unavailable URLs had the 404 error message and the .edu was the most persistent domain with 36% active links. Parker (2007) reported that one of the problems which generated the error 404 message was the use of full stops at the end of URLs, meaning that URLs fail because they were typed incorrectly. Falagas *et al.* (2007) explored accessibility of online resources of Lancet and New England Journal of Medicine and found that 62.2% of online resources were inaccessible. Wagner's *et al.* (2009) study on accessibility analysis of 2011 unique URLs from five dominant journals in medical healthcare management journals from 2002 to 2004 showed that 49.3% of URLs were unavailable. Isfandiari and Saberi (2010) examined the accessibility and half-life of cited URLs in the published papers in the *Information Research Journal* from April 1995 to March 2008. The study found that *.org* and *.net* domains had the most stability and 73% of the URLs were accessible. In tracking web citations in research papers of undergraduate students, Davis and

Cohan (2001) reported that 45% of the cited URLs had disappeared after 12 months and 82% of URLs failed after three years.

Research has also established that some web resources may persist longer than others; meaning that different online resources have different half-lives. For instance, Kumar and Kumar (2012) have cited many studies (Koehler, 2002; Rumsey, 2002; Markwell and Brooks, 2003; Spinellis, 2003; Bar-Ilan and Pertiz, 2004; Sellitto, 2005; Goh and Ng, 2007; Moghaddam and Saberi, 2010) which show different half-life periods ranging from 1.6 to about 15 years. Rumsey (2002) found that the half-life value of web citations in 500 law review articles from 1997 – 2001 was as low as 1.6 years. Similarly, Koehler (2002) estimated half-life of 361 web pages downloaded between 1999 and 2001 to be approximately two years. Other studies have established slightly higher half-life periods. In examining 515 web pages used in graduate-level biochemistry and molecular biology courses, Markwell and Brooks (2003) obtained a half-life value of 4.6 years. In a study conducted by Goh and Ng (2007), half-life of articles spanning a period of seven years (1997-2003) in three leading information science journals was found to be approximately five years. Recent studies have found much higher values of half-life periods of web citations. For example, Moghaddam and Saberi (2010) found that the average half-life for cited internet resources in the *Information Research* journal was 14.94 years. In a more recent study, Kumar and Kumar (2012) established that the half-life period of web citations in two LIS open access journals approximately 11.5 years.

Generally, research on persistence and decay of cited URLs in academic publications is growing. Many of these studies have focused on the availability and persistence of cited URLs, error messages of decayed URLs, domain types of decayed URLs, and the half-life of the web citations. However, there is scarce literature on accessibility analysis of cited URLs in students' works as most studies have focused on scholarly journals. This study was intended to fill this gap.

STUDY CONTEXT

The Sokoine National Agricultural Library (SNAL) was established by an Act of Parliament No. 21 of 1991, which elevated the former library for the Sokoine University of Agriculture (SUA) to a national agricultural library. SNAL therefore serves both as a university library for SUA as well as a national agricultural library for Tanzania. The library is located at the SUA's main campus and has a branch library at the SUA's Solomon Mahlangu Campus in Morogoro. SUA was established in 1984 out of the former Faculty of Agriculture, Forestry and Veterinary Science of the University of Dar es Salaam. It is the second oldest public university in Tanzania, and until recently it was the only agricultural university in the country. During this study, SUA had a total of 28 undergraduate, 46 masters and five non-degree (certificate and diploma level) programmes mainly in agricultural sciences, forestry, animal sciences, education, food sciences, rural development and information sciences. The University also provides doctorate studies in many of these disciplines.

METHODS

This was a bibliometrics study which employed citation analysis technique to examine the persistence and decay of URLs cited in theses and dissertations dated between 2007 and 2011 that were available at SNAL. Citation analysis is a well-known technique for studying relationships and patterns between citing and cited documents (Olatokun and Makinde, 2009). The study was conducted between September and November 2012 involving a sample of 83 theses and dissertations. In total there were 835 doctoral theses and masters dissertations in the library covering the specified period. These were stratified according to their years of publication in order

to randomly draw a 10 percent sample size from each year. All URLs were manually tested for their accessibility by carefully typing the addresses on browsers' address bar. A URL was considered active if the resource was found at its original location or if redirected to a new location. If the resource was not found, the inaccessible URL was considered as "decayed" and the associated error and domain endings were recorded. Since some sites might have been only temporarily unavailable, inactive links were rechecked several times within seven days and if they were still inaccessible at that time, they were recorded as inactive. If access to a resource returned error message *401 Unauthorized,* the resource was considered available because this had to do with user authentication. It should be noted that since the objective of the study was to examine the persistence and decay of URLs, locating the references did not involve the use of search engines. In order to estimate half-life of web resources cited in theses and dissertations, the procedure used by Moghaddam *et al.* (2010) was employed. Half-life t(h) of web citations for each year was calculated using the formula: t(h) = [*t*ln(0.50)]/ [lnW(*t*)-W(0)], where W(0) is the number of web citations at the time of publication and W(t) is the number of active web citations at some later time *t*. The collected data were analysed using Microsoft Excel.

RESULTS AND DISCUSSIONS

All 83 theses and dissertations had a total of 15,468 citations with an average of 186.4 citations per thesis or dissertation. There were 1,487 (9.6%) web citations for all theses/dissertations with an average of 17.9 web citations per thesis/dissertation. The lowest number of web citations was recorded in 2008 (178 citations) and the highest number was recorded in 2011 (522 citations), suggesting increased use of web resources in theses and dissertation. Of 1,487 web citations, only 625 (42%) were accessible and the rest (58%) were inactive. Surprisingly, web citations in newer theses and dissertations were less accessible than those in older theses and dissertations. For instance, the most accessible web citations were recorded for the year 2007 (47.1%) and the most decayed web citations were recorded for the year 2011 (59.6%) (Table 2). This suggests that even newly cited web resources can disappear rather quickly. Some previous studies such as Davis and Cohan (2001), Falagas *et al.* (2007) and Wagner's *et al*. (2009) had found even higher decay rates.

Table 2: *Distribution of citations in theses and dissertations*

Year	Available theses and dissertations	Theses and dissertations sampled	Total citations	Average citation	Web citations	Average Web citations	Active Web citations (%)	Inactive Web citations
2007	156	16	2,121	132.6	187	11.7	88 (47.1)	99 (52.9)
2008	154	15	2,714	180.9	178	11.9	64 (36.0)	114 (64.0)
2009	243	24	2,707	112.8	185	7.7	90 (46.8)	95 (51.4)
2010	145	14	2,904	207.4	415	29.6	172 (41.4)	243 (58.6)
2011	137	14	5,022	358.7	522	37.3	211 (40.4)	311 (59.6)
Total	**835**	**83**	**15,468**	**186.4**	**1,487 (9.6)**	**17.9**	**625 (42.0)**	**862 (58.0)**

When error messages associated with decayed URLs were recorded, the *404 File Not Found* error with 799 messages (92.7%) was the most encountered. This means that web citations in the theses and dissertations had disappeared from their original locations and nothing was matching the requested URLs. However, this figure is much higher than those reported in previous studies (Casserly and Bird, 2003; Goh and Ng, 2007). The *404 File Not Found* error is largely caused by changes in the URL such as the removal or relocation of files as well as changes in file or directory names. Other recorded error messages were *503 Service Unavailable* (3.3%), *403*

Forbidden (1.7%), *410 Gone* (0.9%), *406 Not Acceptable* (0.9%), and *500 Internal Server Error* (0.6%) (Table 3). These findings reinforce the transient and volatile nature of the web as a publishing medium where information resources can be easily altered or removed from their original locations for various reasons including failure or removal of equipment such as servers as well as reconstruction of websites. Errors in typing URLs can also contribute to URL decay.

Table 3: *Type of errors encountered*

Type of errors	Frequency	Percent
404 Not Found	799	92.7
503 Service Unavailable	22	3.3
403 Forbidden	15	1.7
406 Not Acceptable	8	0.9
410 Gone	7	0.9
500 Internal Server Error	5	0.6
Total	**862**	**100**

The findings in Table 4 indicate that of the 861 decayed web citations, the top-level domain *.com* had the greatest number (28.2%) of missing URLs followed by *.org* (18.9%). Nevertheless, the findings indicate that there is little loss associated with country endings (6.5%), government sites (6.6%) and other top level domains (6.7%). These findings support earlier studies (Dimitrova and Bugeja, 2007; Goh and Ng, 2007; Isfandiari and Saberi, 2010) which reported that *.org, .edu* and *.net* were the most persistent domains. The findings are also in line with those of Moghaddam *et al.* (2010) who reported that the *.com* domain was among those with poorer stability and persistence.

Table 4: *Top-level domains associated with decayed URLs*

Year	Decayed citations	.com	.org	.ac	.edu	.gov/ .go	Country endings	.net	Others
2007	99	21	26	11	16	3	8	9	5
2008	114	27	26	14	12	10	8	9	8
2009	94	11	20	17	18	12	3	10	3
2010	243	76	44	18	34	11	10	27	23
2011	311	108	47	40	26	21	27	23	19
Total	861	243 (28.2)	163 (18.9)	100 (11.6)	106 (12.3)	57 (6.6)	56 (6.5)	78 (9.1)	58 (6.7)

The procedure used by Moghaddam *et al.* (2010) was employed to estimate half-life of web resources cited in theses and dissertations. The results of this estimation presented in Table 5 indicate that the average half-life for the cited URLs in theses and dissertations was 2.5 years. This means that it takes only two years and six months for half of the web citations in the theses and dissertations to disappear. This average half-life is slightly higher than that reported by Koehler (2002) (2 years) and Rumsey (2002) (1.6) years. However, higher figures had been

reported by Markwell and Brooks (2003) (4.6 years), Goh and Ng (2007) (5 years), Kumar and Kumar (2012) (11.5 years) and Moghaddam and Saberi (2010) (14.94 years). Generally, these findings emphasize the ephemeral nature of the web resources that makes relevant information disappear.

Table 5: Half-life of Web citations

Year	Half-life (years)
2007	4.6
2008	2.7
2009	2.9
2010	1.6
2011	0.8
Average	**2.5**

Generally, the findings in the present study have shown that many web resources cited in theses and dissertations available at SNAL have disappeared from their original locations. This is contrary to the desire of many citation conventions, styles and guidelines that require authors to cite URLs as part of their bibliographic details so as to ensure access to the cited web resources. This disappearance of web resources has important implications for the scholarly community. It might prevent readers from going backwards to consult primary sources that supported new ideas. This in turn contradicts the principle that the advancement of knowledge is incremental with authors *standing on the shoulders of others in furthering their own work* (Sellitto, 2004). Missing the cited resources hampers readers from seeing the way previous authors *provided the shoulders* for other authors to support their arguments, substantiate claims, build upon existing works, provide contexts of their researches, and compare approaches and methodologies. The study has also shown that URLs with certain top-level domains such as *.com* have higher levels of decay than others.

CONCLUSION

The results of this study indicate that many web resources cited in the doctoral theses and masters dissertations available at SNAL have disappeared from their original locations. The most common error message was *404 File Not Found* and the *.com* top-level domain had the highest number of missing URLs. The results have also shown that it takes only two years and six months for half of the web citations in these theses and dissertations to disappear. This lack of persistence of web references implies that the long term availability of online information resources cannot be guaranteed which in turn raises questions as to whether URLs should continue to be included as part of bibliographic details in the lists of references. Efforts are therefore required from various actors including authors, editors, publishers, libraries, web managers and ICT professionals in order to reduce the problem URL decay. For example, while authors are argued to be careful when typing the URLs, editors need to become more pro-active in their roles as quality controllers. This study supports a number of recommendations previously made by other authors aiming at increasing the availability of URLs. These include the need for authors to retain digital backup or printed copies of cited web resources; advocating for the inclusion of web content in Internet archives; checking URLs systematically before publishing; using Digital Object Identifiers (DOIs) and Uniform Resource Names (URNs) in place of URLs; and establishing institutional repositories in order to upload copies of scholarly material such as

preprints. In addition, the use of citation and referencing software would avoid citation errors resulting from typing. Further research could be conducted to retrieve missing URLs using other tools such as search engines. Furthermore, a comparative study is required to investigate the accessibility and decay of web citations in different types of publications such as journals and theses.

REFERENCES

Bar-Ilan, J. & Peritz, B. (2004)." Evolution, continuity, and disappearance of documents on a specific topic of the web: A longitudinal study of informetrics" J*ournal of the American Society for Information Science and Technology*, 55(11): pp. 980-990.

Casserly, M. and Bird, J.E. (2003). "Web citation availability: analysis and implications for Citations". *American Communication Journal,* 9(2). Retrieved 11th January 2013 from http://crl.acrl.org/content/64/4/300.full.pdf+html

Davis P. M. and Cohen S. A. (2001). "The Effect of the Web on Undergaduate Citation Behavior 1996-1999." *Journal of the American Society for Information Science and Technology*, 52 (4): pp. 309-314.

Decay of uniform resource locators in health care management journals". *Journal of the Dermatology,* 142(9): pp.1147-1152. Retrieved 9th February 2013 from http://archderm.amaassn. Org/cgi/reprint/142/9 /1147.pdf

Dimitrova, D.V. and Bugeja M. (2007). "Raising the dead: recovery of decayed online Citations". *American Communication Journal,* 9(2). Retrieved 11st January 2012 from http://www.quosafulltext.com/sc_ddm/sc_ddm.jsp

Falagas, M.E. Karveli, E.A. and Tritsaroli, V.I. (2007). "The risk of using the Internet as reference resource: A comparative study". *International Journal of Medical Informatics*, 77(4):pp. 280-286. Retrieved 27th December 2012 from http://www.ncbi.nlm.nih.gov/pubmed/17714983

Germaine C. A. (2000). "URLs: Uniform Resource Locators or Unreliable Resource Locators?" *College & Research Libraries*, 61 (4): pp. 359-365.

Goh, D.H.and Ng, P.K. 2007."Link decay in leading information science journals". *Journal of the American Society for Information Science and Technology,* 58(1): pp. 15-24. Retrieved 17th February 2013 from http://onlinelibrary.wiley.com/doi/10.1002/asi.20513/pdf

Harter, S. and Kim, H. (1996). Electronic journals and scholarly communication: a citation and reference study. *Information Research* 2 (1) paper 9. Retrieved 27th December 2012 from http://informationr.net/ir/2-1/paper9a.html.

Koehler, W. (2002). "Web page change and persistence - a four-year longitudinal study". *Journal of the American Society for Information Science and Technology*, **53**(2): pp. 162-171.

Koehler W. (1999), "An Analysis of Web Page and Web Site Constancy and Permanence." *Journal of the American Society for Information Science, 50 (2): pp. 162-180*. Retrieved 25th January 2013 from http://onlinelibrary.wiley.com/doi/10.1002/%28SICI%291097-4571%281999%2950:2%3C162::AID-ASI7%3E3.0.CO;2-B/abstract

Krause, S. (2007). "The process of research writing: citing your research using MLA or APA style". Retrieved 22nd January 2013 from http://www.stevendkrause.com/tprw/

Kumar, B.T. S. and Kumar, K.S. M. (2012). Persistence and half-life of URL citations cited in LIS open access journals. *Aslib Proceedings*, 64 (4): 405 – 422.

Lawrence, S. Coetzee, F. Glover, E. Pennock ,D. M. Flake, G. and Nielsen, F. (2001). "Persistence of Web References in Scientific Research*." IEEE Computer, 34 (2): pp. 26-31.*

Markwell J. and Brooks D. W. (2002). "Broken Links: The Ephemeral Nature of Educational WWW Hyperlinks." *Journal of Science Education and Technology*, 11: pp. 105-108.

Markwell J. and Brooks D. W. (2003). "Link Rot Limits the Usefulness of Web-based Educational Material in Biochemistry and Molecular Biology." *Biochemistry and Molecular Biology Education*, 31: pp. 69-72. Retrieved 16th February 2013 from http://onlinelibrary.wiley.com/doi/10.1002/bmb.2003.494031010165/abstract

Moghaddam, A. and 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): pp. 57-75. Retrieved 16th February 2013 from http://www.webpages.uidaho.edu/~mbolin/moghaddam-saberi.htm.

Parker, A. (2007). "Link Rot: How the Inaccessibility of Electronic Citations Affects the Quality of New Zealand Scholarly Literature". *Library Journal Articles.* Retrieved 5th June 2013 from *http://repository.digitalnz.org/system/uploads/record/attachment/215/link_rot__how_the_i naccessibility_of_electronic_citations_affects_the_quality_of_new_zealand_scholarly_lite rature.pdf*

Powell T. (2003). *HTML & XHTML: The Complete Reference,* 4th edition. Berkeley: Osborne/ McGrawHill.

Rumsey M. (2002). "Runaway Train: Problems of Permanence, Accessibility, and Sustainability in the use of Web Sources in Law Review Citations." *Law Library Journal*, 94: pp. 27-39.

Sellito, C. (2004) , **"**A study of missing Web-cites in scholarly articles: towards an evaluation framework". Journal of information science, 30(6): pp.484-495. Retrieved 16th November 2012 from   http://jis.sagepub.com/content/30/6/484.refs

Spinellis D. (2003). "The Decay and Failures of Web References." *ACM*, 46 (1): pp. 71-77.

Tajeddini, O. Azimi, A. Sadatmoosavi, A.and Sharif-Moghaddam, H. (2011)."Death of web citations: A serious alarm for authors". *Malaysian Journal of Library & Information Science, 16(3):pp.17-29.*

Ticehurst G. W. and Veal A. J. (2000). *Business Research Methods,* Frenchs Forest: Longman.

Tyler, D., & McNeil, B. (2003). "Librarians and link rot: A comparative analysis with some methodological considerations". *Portal: Libraries and the Academy, 3* (4): pp. 615–632. Retrieved 3rd September 2012 from http://digitalcommons.unl.edu/libraryscience/62.

Wagner, C. Gebremichael, M.D. Taylor, M.K. and Soltys, M.J. (2009). "Disappearing act: Decay of uniform resource locators in health care management journals". *Journal of the Medical Library Association,* 97(2): pp.122-130. Retrieved 21st December 2012 from http://www.ncbi.nlm. nih.gov/pmc/articles/PMC2670212/?tool =pubmed.

Webster J. and Watson R. T. (2002). "Analyzing the Past to Prepare the Future: Writing a Literature Review." *MIS Quarterly*, 26 (2): pp. 13-23.

Wren, J.D. (2008). "URL decays in MEDLINE a 4-year follow-up study". *Bioinformatics*, 24(11):1381-1385. Retrieved 15th November 2012 from http://bioinformatics.oxfordjournals.org/content/24/11/1381.full

Wren, J.D. Johnson, K.R. Crockett, D.M. Heilig, L.F. Schilling, L.M. and Dellavalle, R.P. (2006). "Uniform resource locator decay in dermatology journals". *Archives of Dermatology,* 142(9): pp.1147-1152. Retrieved 9th January 2013 from: http://archderm.amaassn. Org/cgi/reprint/142/9 /1147.pdf

Original article at: http://ijedict.dec.uwi.edu//viewarticle.php?id=1606