Library Progress International
Vol.43, No.2, July-December 2023: P.236-243
DOI: 10.48165/bpas.2023.43.2.11

Print version ISSN 0970 1052
Online version ISSN 2320 317X

Original Article

Available online at www.bpasjournals.com

An Analysis of the Availability and Persistence of Web Citations in Journal of Indian Library Association (JILA)

[1]Manjuli Devi* and [2]Aman Verma

Author's Affiliation:
[1]Assistant Librarian, Innovative Group of Colleges, Greater Noida, Uttar Pradesh 201310, India
E-mail: dixitmanjuli95@gmail.com
[2]Senior Research Fellow, Department of Library and Information Science, University of Lucknow, Lucknow, Uttar Pradesh 226007, India
E-mail: vermaaman.1443@gmail.com

***Corresponding Author:** Manjuli Devi, Assistant Librarian, Innovative Group of Colleges, Greater Noida, Uttar Pradesh 201310, India
E-mail: dixitmanjuli95@gmail.com

(Received on 26.07.2023, Revised on 28.09.2023, Approved on 11.10.2023, Accepted on 11.10.2023, Published on 21.12.2023)

How to cite this article: Devi M. and Verma A. (2023). An Analysis of the Availability and Persistence of Web Citations in Journal of Indian Library Association (JILA).*Library Progress International,* 43(2), 236-243.

ABSTRACT
The present study aims to investigate the uses of URLs as citations based on 148 articles published in the Journal of the Indian Library Association (JILA) published during 2018-2021.Data has been collected from the websites of the journal. Data collection is carried out by visiting selected academic journal websites.A total of 2319 citations were retrieved and analyzed, and it is observed that there were 1373print citations and 946 web citations. Among 946 web citations, 751are active while195 are missing web citations. It has been found that HTTP 404 error codes accounted for more thanhalf the life of any error code associated with web citations. The average half-life for the Journal of Indian Library Association is 8.818 years. This research provides unique insights into the prevalence and characteristics of URL citations within the specified journal.

KEYWORDS: *Web citations, Journal of Indian Library Association, Indian Library Association, Citation Analysis, Half-Life.*

INTRODUCTION

A reference is an acknowledgment that one document guise to another,orareference is a detailed entry that provides all the necessary information about a source cited in the textbook. References generally appear at the end of a scholarly paper, composition, or explorationreport in a section generally appertained to as the References Works Cited or Bibliography depending on the citation style being used for example- APA, MLA, Chicago, etc. The purpose of furnishing references is to give credit to the original authors or generators of the workshop cited and enable compendiums to detect those sources for further reading or verification of information. References are an essential element of academic integrity, as they demonstrate the depth of exploration and the use of believable and dependable sources to support the arguments or findings presented in a scholarly paper. A citation is a way you tell

your compendiums that certain material in your work came from another source. It also gives compendiums of the information necessary to find the position details of that source. In general, a citation implies a relationship between a part or the total of the cited document and a part or the total of the citing document.

A web citation, also known as an online citation or electronic citation, is a specific type of reference that provides the necessary information to identify and locate online sources used in academic writing, research, or any other form of digital content. Web citations are used when referencing websites, online articles, web pages, blog posts, online reports, and other electronic resources available on the internet. It includes the author's name, the title of the web page or article, the URL, the publication date, and any other relevant information required by the citation style being used. Web citations are essential for giving credit to the sources and enabling readers to access and verify the referenced online material easily.

The half-life" of a web citation refers to the time it takes for a web link or reference to become inactive or no longer accessible. In other words, it is the average time it takes for a URL or web page to become outdated or unavailable. In this study, the author tries to analyze the longevity of the web citations available in the selected journal.

REVIEW OF LITERATURE

With the emergence of the internet due to its ability to access data quickly and conveniently,thestudy and use of web citations have been considered by scholarly researchers. Thus, many researchers have been working on this topic and could document some related issues in this field. Here, some related studies that considered the behavioral effects of web citations on scholarly communication have been explained.

Goh and Ng (2007) conducted a study on Library and information journals during 1997-2003. Only 69 % of those URLs were permanent, while the remaining 31 percent had disappeared from the original web address. 56% of error messages were "404" (page not found). The ".edu" with 36% active links was the most table domain.

Casserly and Bird (2008) examined a collection of five hundred citations that referenced materials from articles published in journals related to Library and information science between the years 1999 and 2000that were scrutinized and probed on the internet. A significant proportion of these citations provided only partial bibliographic details and lacked any indication of the date they were accessed. The majority of URLs led to pages containing educational ("Edu") or organizational ("org") content and did not incorporate a tilde (~) in their structure. A noteworthy observation was that over half of the URLs (56.4%) demonstrated permanence, and around 81.4% remained accessible on the web. It was identified that attributes such as content, domain, and directory depth played a role in determining the accessibility of the resources. The study concludes by offering eight recommendations aimed at enhancing citation conventions for scholarly communication.

Kumar and Kumar (2017) conducted a study on the utilization of URLs as citations within scholarly publications in the field of Library and Information Science. They examined a total of 8203 research articles published across 12 LIS journals during the years 2006-2015. Out of the 288,452 citations analyzed, a substantial 42,098 comprised URLs. The researchers also delved into the attributes linked to the cited URLs. The findings of the study indicated that single-authored papers averaged 5.42 URL citations per article. Moreover, the research highlighted the significant citation of URLs associated with both organizational and commercial domains. Notably, HTML and PDF file formats emerged as dominant in the citations.

Wu (2009) examined the accessibility of 1,637 web references found in two notable Chinese academic journals, encompassing the years between 1999 and 2003. The findings indicated that web references per article grew by 26.2%, surpassing the growth rate of conventional

references, which was at 43.6%. It is essential to highlight that the study found that only 44.2%, equivalent to 723 web-references, were accessible.

McCown (2001) conducted a study availability and persistence of URLs cited in articles published in D-Lib Magazine. For doing their research, they extracted 4387 unique URLs referenced in 453 articles published from July 1995 to August 2004. Inconclusion, it was realized that approximately 28% of those URLs failed to resolve initially, and 30% failed to resolve at the last check. Most of the unresolved URLs were due to 404 (page not found) and 500 (internal server error) errors. Moreover, based onthe data collected, they found the half-life of a URL referenced in a D-Lib Magazine articleis approximately 10 years. It was also found that URLs were more likely to be unavailableif they pointed to resources in the .net, Edu, or country-specific top-level domain, usednonstandard ports (i.e., not port 80), or pointed to resources with uncommon or deprecatedextensions.

OBJECTIVES

The objectives of the study are:

1. To find out the total number of citations and proportion of print and web citations.
2. To find out the active and missing web citations.
3. To find out the type of HTTP Errors associated with missing web citations.
4. To calculatethe half-life of web citations.

SCOPE & METHODOLOGY

The Indian Library Association is a professional organization that represents the library and information science community in India. It aims to promote growth publications of serial and other publications for dissemination of information, and it is the promotion of research bibliographical studies.In this investigation, an analytical survey is the chosen methodology, and online sources are utilized for obtaining articles.Thedata was collected from the websites of the selected journals, where each article was manually retrieved, and the corresponding citations were meticulously documented.The present study aims to analyze the web citations of the journal, The Journal of IndianLibrary Association during the period 2018-2021. The total number of citations found in the study was 2,319. Each URL was individually checked, and those that successfully opened were categorized as "active," while those that failed to open were categorized as "missing." The focus of the study is to find out the half-life of web citations.In modern publications, it is generally easier to access web references than in older papers. Consequently, web citations tend to become less accessible as time elapses following the publication of articles. Thus, the author has chosen to focus on web citations for the period 2018-2021.

DATA ANALYSIS

Table 1: The proportion of URLs used as citations

Journal	Number of articles	Total citations	Print citations (%)	Web citations
The Journal of the Indian Library Association	148	2319	1373(59.20%)	946(40.79%)

Table 1 shows an overview of the citation count of The Journal of the Indian Library Association for the period 2018- 2022. The findings reveal that a total of 148 articles were published, accumulating 2319 citations, out of which 1373 were print citations and 946 were web citations.

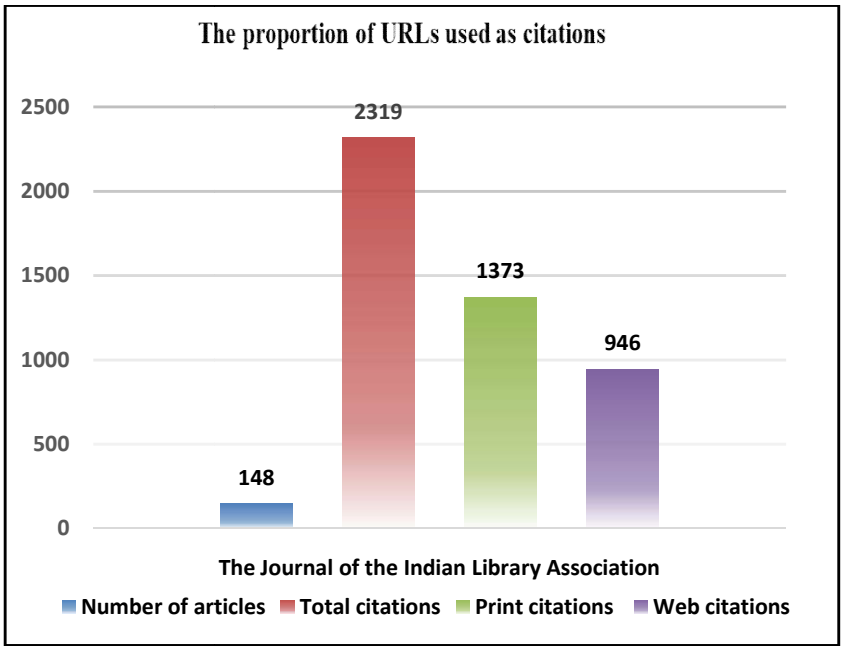

Figure 1: The proportion of URLs used as citations

Table 2: The proportion of URLs used as citations cross-tabulated by year

Year	Number of articles	Number of articles with URLs (%)	Total citations	Print citations (%)	Webcitations(%)
2018	23	16(69.56%)	311	158(50.80%)	153(49.19%)
2019	28	22(78.57%)	384	265(69.01%)	119(30.98%)
2020	41	32(78.04%)	636	430(67.61%)	206(32.38%)
2021	56	52(92.85%)	988	520(52.63%)	468(47.36%)
Total	**148**	**122(66.30%)**	**2319**	**1373(59.20%)**	**946(68.90%)**

Table 2 presents the count of URLs used as references in the Journal of Indian Library Association. In 2018, a total of 311 citations were recorded; in 2019, the count increased to 384; and in the years 2020 and 2021, the counts reached 636 and 988 citations, respectively. Among the 311 citations in 2018, 96% (153) were web citations. In 2019, out of the 384 citations, 44% (119) were web citations. Overall, across 148 published articles, there were 68% (946) web citations out of a combined 2319 citations.

Table 3: Proportion ofactive and missing web citations

Journal	Total web citations	Active web citations (%)	Missing web citations (%)
The Journal of the Indian Library Association	946	751(79.38%)	195(20.94%)

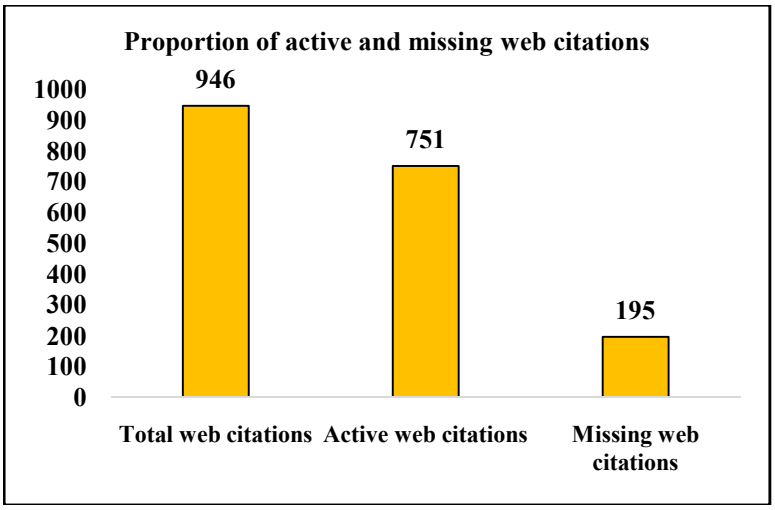

Figure 2: Proportion of active and missing web citations

Table 3 shows the percentage of active and missing web citations in the journal. A total of 946 citations were found out of which about 79 percent of total web citations were still active and the remaining were missing.

Table 4: Year-wise distribution of missing web citations

Year	Total web citations	Active (%)	Missing (%)
2018	153	108(70.58)	45(29.41)
2019	119	108(90.75)	11(9.24)
2020	206	160(77.66)	46(47.57)
2021	468	375(80.00)	93(19.87)
Total	**946**	**751(79.38)**	**195(26.10)**

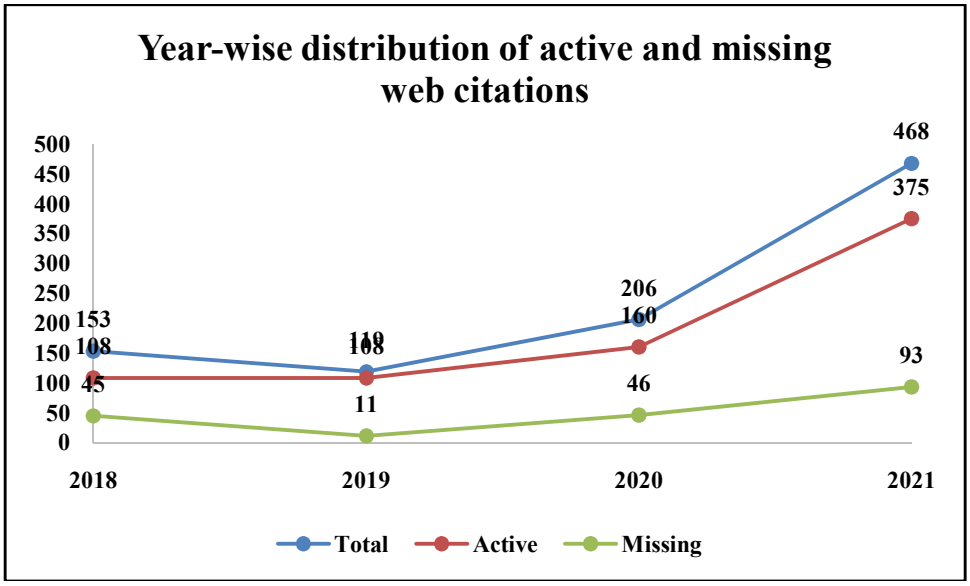

Figure 3: Year-wise distribution of active and missing web citations

Table 4 presents the year-wise distribution of web citations in the Journal of the Indian Library Association. It reveals that in the year 2018, there were 108 active web citations out of 153, indicating that only 29% of web citations are missing. In the year 2019, there were 108 active web citations, succeeded by 160 and 375 active web citations in the years 2020 and 2022, respectively. The year 2021 sees a mere 19.87% of web citations missing.

Table 5: HTTP errors associated with web citations

HTTPErrors	Name of errors	Total
301	Moved permanent	1
303	Moved temporarily	1
400	False request	1
401	Unauthorized response	1
403	forbidden	1
404	File not found	175
502	Invalid response	2
503	Service Unavailable	4
Can't reached	-	9
Total		**195**

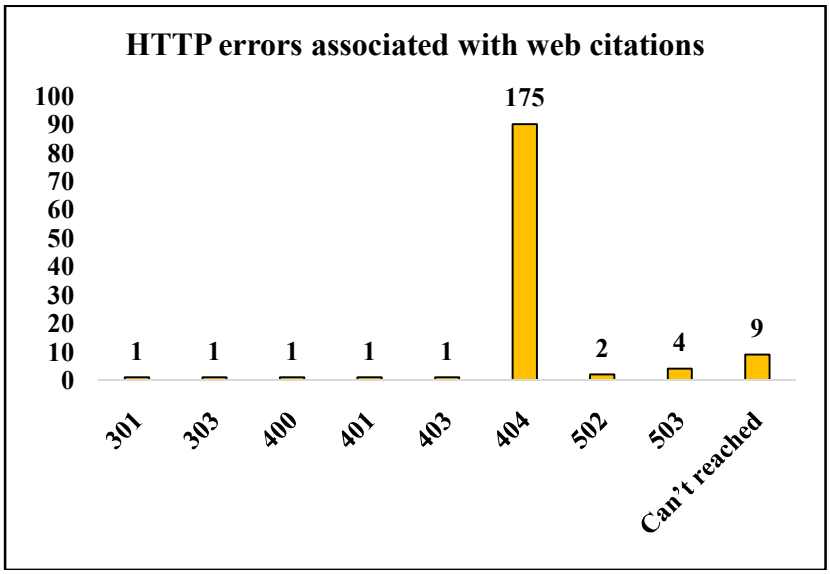

Figure 4: HTTP errors associated with web citations

HTTP status codes of missing web citations are presented in Table 5. The HTTP 404 error code constitutes the majority of error instances linked to web citations. It signifies that the server cannot find the requested resource. It is clear from the table that there were 175 errors associated with error code HTTP 404, which is about half of any error code associated with web citations. Besides these typical HTTP errors, there are nine other error instances categorized as 'Unreachable.'

Half-life estimation of web citations

Half–life is the period required for half of a defined web citation to disappear.Half-life is the

period required for half of a defined web citation to disappear. Therefore, to make a clear estimation of the half-life of web citations for each year, the following formula is used to calculate the half–life of online citations.

t_h= [t In (0.5)] / [In $w(t)$ - In $w(0)$]

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

Table 6: Calculation of Half-life of web citations

Year	**Time (t)**	**Total number of web citations w(o)**	**Total number of active citations w(t)**	**Half-Life**
2018	4	153	108	7.963
2019	3	119	108	21.44
2020	2	206	160	2.742
2021	1	468	375	3.128
Average Half-Life				**8.818**

Table 6 depicts the calculation of the half-life of web citations along with the average half-life. The mean half-life for the Journal of Indian Library Association and Information is calculated at 8.818 which indicates on average, after 8 years, approximately half of the web citations or links will have become inaccessible or broken. In other words, if you were to track a collection of web citations over time, you would expect that about 50% of them would no longer lead to valid web content after this 8-year period.

This estimate is a way to illustrate the decreasing reliability and accessibility of web citations over time due to various factors such as changes in web content, website restructuring, link rot, and the dynamic nature of the internet. It emphasizes the importance of regularly checking and updating citations in academic and research work to ensure the accuracy and validity of references. Additionally, it underscores the value of using permanent identifiers like DOI (Digital Object Identifiers) when available, as they are designed to remain stable and provide a reliable way to access scholarly content.

CONCLUSION

The increasing use of URLs as citations reflects the evolving nature of research practices, where digital resources and online content play a significant role in scholarly communication.It is observed that the overall percentage of web references in journal articles is slowly increasing, but they still do not match the prevalence of print references, reflecting a common trend in the academic publishing world. Over time, websites can undergo changes, including reorganization, renaming, or discontinuation. When the URLs cited in the journal articles are affected by these changes and are not properly maintained or updated, it can lead to link rot, making the citations inaccessible. Supporting the development and maintenance of the Internet Archive (commonly known as the Wayback Machine) and encouraging authors to upload cited URLs is indeed a valuable strategy to address the issue of link rot and ensure the long-term accessibility of web citations.Researchers and authors are encouraged to take proactive steps, such as periodically checking and updating web citations, using permanent identifiers like DOIs when available, and utilizing web archiving services to capture and preserve online content. Additionally, journals and publishers may implement guidelines or requirements to ensure the reliability of web citations in scholarly publications.

REFERENCES

1. Casserly, M. F., & Bird, J. E. (2008). Web citation availability: A follow-up study.

2. Goh D H L and Ng P K, (2007). Link decay in leading information science journals. Journal of the American Society for Information Science and Technology. 58(1), 15-24.

3. McCown, F., Chan, S., Nelson, M. L., &Bollen, J. (2005). The availability and persistence of web references in D-Lib Magazine. arXiv preprint cs/0511077.

4. Vinay Kumar D and Sampath Kumar B T. (2017). Prevalence of URLs in Library and Information Science (LIS) Literature: A Citation Analysis. collect Journal of Scientometrics and Information Management. 11(2), 287-297.

5. Wu, Z. (2009). An empirical study of the accessibility of web references in two Chinese academic journals. Scientometrics. 78(3), 481-503.

****