Web Citations in Indian Theses: Exploring the Use of Web Citations by Social Scientists

B T Sampath Kumar
Professor and Chairman, Project Director, Major Research Project, ICSSR
Department of Library and Information Science, Tumkur University
Tumakuru, Karnataka, INDIA
sampathbt2001@gmail.com

Dhanaraj
Research Assistant, Major Research Project
ICSSR, Department of Library and Information Science
Tumkur University, Tumkur, Karnataka, INIDA
dhanarajslokane9108@gmail.com

ABSTRACT: *Web sources offer numerous advantages for academic research and allow researchers to connect with other researchers. This study explores the percentage of web sources cited in the selected Indian Social Science theses. Which web citation domain was the most frequently used by Indian social scientists? Further, the study also identified which file format of web citations is commonly used by social scientists. For this purpose, 433 theses in the selected five subjects were downloaded from the Shodhganga repository, and the Web references cited in the theses were extracted. It was found that 20.45% of references were web citations. A higher number of web citations are found in Journalism and Mass Communication theses (30.91%), followed by Sociology (24.52%) and Library and Information Science (30.70%). The most popular domain was.com/.co among the Social Sciences Scientists 2631 (22.01%). The more significant number of web citations have a path depth of 2 (50.94%).*

Keywords: Web Citations, Indian Theses, Social Scientists, and References

Received: 22 July 2024, Revised 28 July 2024, Accepted 10 August 2024

DOI: https://doi.org/10.6025/stm/2024/5/39-50

1. Introduction

The Internet has become a crucial reference tool for scientists and the general public to access timely and valuable information because scientific knowledge and information are created digitally (Bansal, 2021). Due to its rapid expansion, the Internet is becoming a commonly used research and information resource. In recent years, much information has been shared and exchanged via the Internet. Since the invention of the Internet, there has been a notable increase in the quantity of scholarly and scientific e-resources, including e-books, e-theses and dissertations, e-prints of research articles, e-journals, etc. Citing online resources in academic papers has become more common as e-resources are more readily available online (Sampath Kumar & Prithviraj, 2015).

However, the dynamic nature of web sources is one of their issues since web resources unexpectedly change or vanish over time. There is a lot of concern about the longevity, availability, and suitability of links to web resources. Because the web is a dynamic environment with no assurance of permanence. Thus, academics are increasingly concerned about the accessibility of these sources. (Carnevale and Aronsky, 2007). The process by which a website link or URL gradually loses its functionality over time is referred to as web decay. The rotting web citations are a serious problem for the academic

community. If the cited content is unavailable to readers, the references may lose academic value. Because web publications are extremely fleeting, preserving digital resources is critical for modern civilizations. To tackle the problem of rotting web citations, conserving electronic resources has become more important than accessibility (Niveditha et al., 2021).

In this context, the current study tried to assess the accessibility of web citations in theses published to Shodhganga by NIRF-ranked colleges between 2020 and 2023. Aside from checking for online citations, the script extracts URL characteristics such as file type, top-level domain, path depth, and character length. This data was collected to examine how URL characteristics influence their accessibility.

2. Review of Literature

Over the last couple of years, many studies have attempted to explore the use of web resources in the scholarly literature. This section reviews some of the significant contributions.

In an early work, Prithviraj and Sampath Kumar (2015) analysed the web citations in Indian LIS journals published between 2001 and 2010**.** They found that the mean web citations are 2.60 in the studied sample. Organizational (.org), and commercial (.com) top-level domains had the highest citations (30.91% and 22.08%, respectively). HTML file formats accounted for the most significant percentage of web citations (67.30%), followed by PDF file formats coming in second (11.39%).

Vinay Kumar and Sampath Kumar (2017) studied the URLs as references in academic works in Library and Information Science. Of the 288,452 citations found in the 8203 research articles published in 12 LIS journals between 2006 and 2015, 42,098 were URLs. Likewise, an analysis was conducted on the attributes linked to the mentioned URLs. The study showed 5.42 URL citations on average per piece among the single-authored papers. The analysis also revealed that the HTML and PDF file formats received the most citations, whereas URLs connected to commercial and corporate domains received very few citations.

Vishal's bibliometric analysis of 295 articles published in DESIDOC Journal of Library & Information Technology (DJLIT) between 2011 and 2015, which included 4821 cited documents as a supplement, aimed to determine the authors' utilization of various reference sources and the quantitative distribution of citations in the DJLIT (Vishal, 2017). The results indicated that 1912 (39.65%) citations were attributed to a single author, with two authors coming in second place with 1152 (23.89%), three authors in third place with 456 (9.45%) citations, and more than three authors with 386 (8%) citations. The analysis also showed that journals were the most cited information source, accounting for 2560 (53.10%) of the citations, followed by websites (22.69%), and books (10.81%). Furthermore, the study found that 295 articles were cited across the 30 issues, and the analysis showed a slowdown in the quantitative rise of citations in 2011 and 2015.

In their study, Preeti and Anil Kumar (2017) examined the "Citation Analysis of Doctoral Theses," which comprehensively analyses citation trends in Sociology theses at Panjab University, Chandigarh, between 2002 and 2012. The authors emphasize the preference of researchers for traditional sources, with books constituting 57.55% and journals comprising 36.52% of all citations. Additionally, the study reveals that single authorship is preferred in books (71.24%) and journals (64.97%). The paper received an impressive score of 90%. In the same year, Vinay Kumar and Sampath Kumar (2017) examined 5197 URLs cited in two open-access LIS journals between 2009 and 2013. The author used the W3C link checker to verify the URLs. It was found that 417 out of 5197, or 38.12 per cent) were missing, and the remaining 61.88 per cent of URLs were operational. However, only 30.70 per cent of all disappeared URLs were retrieved by Google; in contrast, 66.19 per cent of all vanished URLs were recovered by the Internet Archive. In the same year, Rupesh Kumar A. (2017) examined a sample of electronic theses in Library and Information Science to determine the prevalence of Web citations. It assessed various attributes, including content type, accessibility, domain composition, path depth, file formats, and URL character length. A total of 2125 Web citations were extracted from 285 ETDs. Journal articles comprised the majority (62.3%) of the Web citations, while theses and dissertations represented the least (3.8%). More than half (51%) of the Web citations had DOIs. The accessibility rate of the Web citations was 85.14%.

Jianhua *et al.* (2018) analyzed 7574 articles published in 10 Information Science journals between 2009 and 2016, encompassing 20,960 references. They found that the co-citation analysis of documents revealed significant shifts in the subjects of the study within the IS domain. Further same year, the study conducted by Amritpal and Gurjeet (2018) revealed that the average number of citations per thesis from 2000 to 2014 was 98.4, based on an analysis of 9053 web citations from 92 doctoral theses uploaded to Punjab University. It was found that journals were the most cited sources, accounting for 49.39 percent of the citations. The author's primary objective was to investigate the citation authorship pattern, the literature format mentioned, and the geographical distribution of citations from books and core journals. The results showed that books have a longer half-life than journals, with a half-life of 21.74 and 18.72 years, respectively. Additionally, the use of electronic resources has become increasingly prevalent.

The study conducted by Manjunatha, *et al.* (2020**)** investigated the usage of URLs as citations and their long-term viability. Based on 966 articles published in specified LIS journals between 2011 and 2015, with an average of 38.26 references per article, the study found a total of 36,968 references are found in the articles. Of these, 5,867 were URL references. The study also discovered that 2,730 of the 5,867 URLs, or 46.53%, were still active, while 3,137 (53.46%) were found to be missing. The highest percentage of broken URLs (83.22%) was reported in LIS papers in 2012, with the most frequent error (40.36%) being the HTTP error 404 ("file not found") error message. The study suggests that authors should ensure the accessibility of URL citations before including them in the article's reference list.

Niveditha and Mallinath Kumbar (2020) conducted a study on the prevalence of web citations in prestigious scholarly journals of Library and Information Science and Communication and Media Studies. The selection criteria for the journals were based on their high impact factor, and they were published between 2008 and 2017. The researchers collected a total of 12,251 articles and extracted 5,55,428 references. After checking the availability of 1,02,718 URLs, they found that the percentage of web citations increased from 2008 to 2017 in both disciplines.

In the same year, Loan and Shah (2020) conducted a study to identify the persistence and decay of uniform resource locators (URLs) associated with Web references. The decaying of Web references is analyzed about their age, domain, technical errors and error codes. For that purpose, the Web references of the *Journal of Informetrics* were selected for analysis and interpretation to fulfill the set objectives. The references of all the scholarly articles, excluding editorials and reviews published in the *Journal of Informetrics* for five years from 2007 to 2011 were recorded in a text file. Later, the URLs were extracted from the articles to verify their accessibility in terms of persistence and decay. The collected data were then transferred into an Excel file and tabulated for further analysis and interpretation using simple statistical techniques. The results showed that of the 7,409 citations retrieved from 221 articles, 358 (4.8%) were Web citations. These Web citations were assessed to find their persistence and decay. The results reveal that 115 (32.12%) Web references were missing or dead. The most common error associated with missing web citations was That 404 pages were not found, contributing 60% of the total missing citations, followed by 400 bad request errors (35.65%). The domain analysis of missing Web citations depicts that most of the missing URLs were associated with the .gov domain (40%), followed by .edu (29.58%) and .com (26.04%).

Ayush Kumar Patel *et al.* (2021**)** investigated the publication trends of the Webology Journal from 2006 to 2020, identifying 295 papers. The study revealed that the greatest number of citations, 273, and the most papers (92) were published in 2020. Iran led the way with 63 documents, followed by India with 50 and the United States with 25.

In recent years, Teixeira da Silva and Nazarovets (2023) have comprehensively reviewed various papers and opined that Web-based references form an integral part of academic culture and literature, as well as scientific communication. Preserving web-based references cited in-text, or in reference, lists of academic papers is essential to ensure the integrity of knowledge preservation. When access or links to web-based resources and references disappear, or if they are dysfunctional, these phenomena are collectively referred to as reference rot. Even though reference rot degrades the integrity of information in academic papers, no industry-wide solutions currently exist. Given its prominence, the study proposes the Internet Archive as a possible solution to address reference rot but also notes its weaknesses and limitations.

3. Research Questions

The study aimed to answer the following research questions:

a) How many theses have been submitted to Shodhganga by the selected Indian universities from 2020 to 2023?

b) What percentage of web citations are cited in the selected Social Science theses?

c) Which web citation domain was the most frequently used by Indian Social Scientists?

d) Which file format of web citations is commonly used by Social Scientists?

e) What are the web citations' character length and path depth?

4. Hypotheses

The following hypotheses were formulated.

H1: The age of the university and the number of theses uploaded to Shodhganga are positively correlated.

H2: There is a significant positive correlation between the year and the number of theses uploaded to Shodhganga from

2020 to 2023.

H3: The number of web citations has increased over a period of time from 2020 to 2023.

5. Scope and Methodology

Investigating how Social Science researchers use web citations in their theses was the goal of the current study. For this reason, the theses submitted by South Indian universities with National Institute of Ranking Framework (NIRF) rankings. NIRF was launched by the Honourable Minister of Human Resource Development, Government of India. This framework outlines a methodology to rank institutions across the country. The parameters broadly cover "Teaching, Learning and Resources," "Research and Professional Practices," "Graduation Outcomes," "Outreach and Inclusivity," and "Perception" (NIRF, 2024). The theses uploaded to Shodhganga by the universities with NIRF rankings between 2020 and 2023 have been taken into consideration. The Shodhganga" is the name coined to denote the digital repository of Indian Electronic Theses and Dissertations set up by the INFLIBNET Centre (Shodhganga, 2024).

Selection of subjects/theses: The Social Science subjects viz., Economics, Journalism and Mass Communication (JMC), Library and Information Science (LIS), Political Science, and Sociology are covered in this study. Only these five subjects were selected since these subjects received a significant number of web citations as compared to other Social Science subjects.

Collection of web citations: It is noticed that between 2020 and 2023, a total of 433 theses in the selected five subjects were uploaded to the Shodhganga repository. After downloading all 433 theses, the citations listed were extracted and the duplicate references were removed. Web citations and DOIs were identified and separated from print references.

6. Analysis of Data

Table. 6.1 No. of Theses uploaded to Shodhganga during the period 2020-2023

Sl. no	Name of Universities	Year of Establishment	NIRF Ranking	No. of Theses uploaded Shodhganga	Percentage
1	Bharathiar University	1982	21	28	6.47
2	University of Kerala	2009	24	14	3.23
3	Alagappa University	1947	30	19	4.39
4	Mahatma Gandhi University	1983	31	09	2.08
5	Osmania University	1918	36	64	14.78
6	Bharatidasan University	1982	41	47	10.85
7	University of Mysore	1916	44	73	16.86
8	University of Madras	1857	50	109	25.17
9	Maduri Kamaraj University	1966	53	34	7.85
10	Periyar University	1997	59	36	8.31
			Total	433	100.00

Table 6.1 presents the distribution of theses by universities. In the present study, 10 universities from South India that qualified for the NIRF ranking in 2023 have been considered. The theses uploaded by these universities to Shodhganga

between 2020–2023 have been downloaded. It was observed that these universities have uploaded a total of 433 theses during the period 2020–2023. The University of Madras has uploaded the highest number of theses (109), whereas Mahatma Gandhi University has uploaded the lowest (09) number of theses on Shodhganga. To know the correlation between the age of the Universities and the number of theses uploaded to Shodhganga, the Correlation has been applied, and it is found that there is a significant positive correlation between the age and the number of theses uploaded to the Shodhganga repository (r=0.879 and p=0.001). Hence hypothesis 1 is accepted.

Table: 6.2 No. of Theses uploaded to Shodhganga by universities (cross-tabulated by subject)

Name of Universities	**Economics**	***JMC**	****LIS**	**Political Science**	**Sociology**	**Total**
Bharathiar University	7	1	16	3	1	28
University of Kerala	8	0	0	6	0	14
Alagappa University	2	0	17	0	0	19
Mahatma Gandhi University	8	0	0	0	1	9
Osmania University 34	0	9	18	3	64	
Bharatidasan University	27	0	18	0	2	47
University of Mysore	33	11	9	16	4	73
University of Madras	87	3	2	7	10	109
Maduri Kamaraj University	6	0	18	3	7	34
Periyar University	11	1	14	2	8	36
Total	223	16	103	55	36	433

**** JMC- Journalism and Mass Communication ** LIS- Library and Information Science***

Table: 6.3 No. of Theses uploaded to Shodhganga by year

Name of Universities	**2020**	**2021**	**2022**	**2023**	**Total**
Bharathiar University	0	2	8	18	28
University of Kerala	0	3	0	11	14
Alagappa University	0	0	11	8	19
Mahatma Gandhi University	0	0	0	9	9
Osmania University	0	1	17	46	64
Bharatidasan University	0	0	27	20	47
University of Mysore	0	24	49	0	73
University of Madras	0	22	41	46	109
Maduri Kamaraj University	0	3	13	18	34
Periyar University	2	9	18	7	36
Total	2	64	184	183	433

Table 6.3 shows the year-wise distribution of theses, cross-tabulated by universities. The highest number of theses were uploaded in 2022 (184) and the lowest in 2020 (02).

Table 6. 4. University-wise distribution of citations and web citations

Name of University	No of theses	No. of citations	No. of print citations	Percen-tage	No. of Web citations	Percentage	Average Web citations per thesis
Bharathiar University	28	3836	2916	76.02	920	23.98	32.86
University of Kerala	14	3573	2218	62.08	1355	37.92	96.79
Alagappa University	19	2627	1662	63.27	965	36.73	50.79
Mahatma Gandhi University	9	1819	1351	74.27	468	25.73	52.00
Osmania University	64	5662	5024	88.73	638	11.27	9.97
Bharathidasan University	47	5672	4624	81.52	1048	18.48	22.30
University of Mysore	73	11211	8601	76.72	2610	23.28	35.75
University of Madras	109	15969	13561	84.92	2408	15.08	22.09
Maduri Kamaraj University	34	3881	3054	78.69	827	21.31	24.32
Periyar University	36	4205	3491	83.02	714	16.98	19.83
Total	433	58455	46502	79.55	11953	20.45	27.61

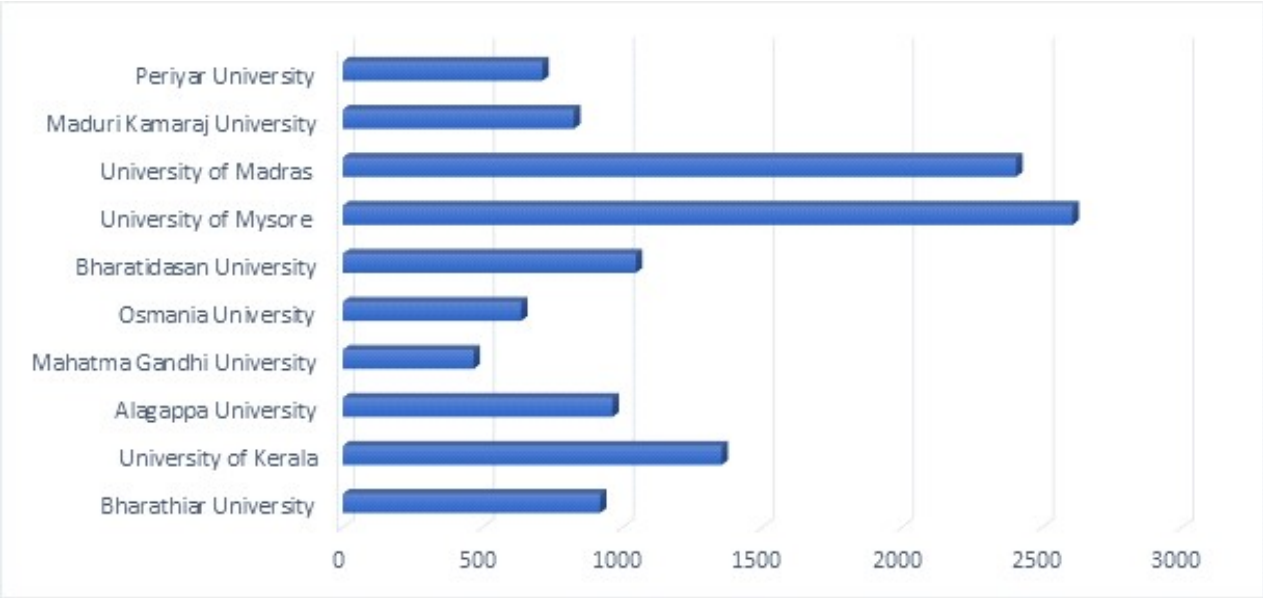

Figure 6.1. University-wise web citations

Table 6.4 displays the distribution of citations and web citations. There were 58455 citations in total, and 46502 print citations, with web citations making up 20.45% of the total. Each thesis has, on average, 135.31 citations. There were 27.61 web citations on average for each thesis. The University of Kerala had the highest average number of citations (255.21), while Osmania University had the lowest average (89.87). With an average of web citations per thesis, the University of Kerala had the most (96.79), while Osmania University had the fewest (9.97).

Table: 6.5. Distribution of citations and Web citations by year

Year	No. of theses	No. of citations	No. of print citations	Percentage	No. of Web citations	Percentage
2020	2	238	222	93.28	16	6.72
2021	64	8856	7470	84.35	1386	15.65
2022	184	24467	19222	78.56	5245	21.44
2023	183	24894	19588	78.69	5306	21.31
Total	433	58455	46502	79.55	11953	20.45

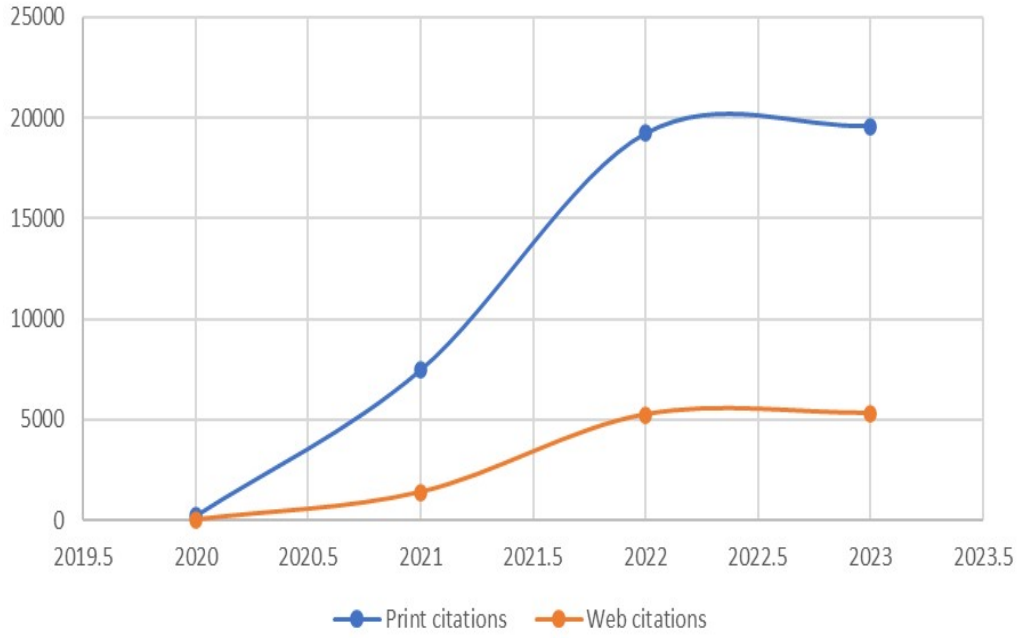

Figure 6.2. Print citations and Web citations by year

The year-wise prevalence of web citations has been presented in Table 6.5. The highest percentage of web citations was found in the year 2023 (21.31%), closely followed by the year 2022 (21.44%). There was the lowest rate of use of web citations in the theses uploaded in the year 2020 (6.72%). A correlation test was conducted to know the type and degree of correlation that exists between the year and the number of theses, citations, and web citations. The results of the test

indicated a high degree of statistically significant positive correlation between year and number of theses (*r = 0.946 p = 0.054*), year and number of citations (*r = 0.952 p = 0.048*), and year and number of web citations (*r = 0.943 p = 0.057*). Hence, there is consistent growth in the number of theses, citations, and web citations. Hence, hypotheses 2 and 3 are accepted.

Table: 6.6. Distribution of citations and Web citations by subject

Subject Percentage	No. of theses	No. of citations	No. of print	citations	Percentage	No. of Web citations
Economics	223	27624	23923	86.60	3701	13.40
*JMC	16	3154	2179	69.09	975	30.91
**LIS	103	12907	8944	69.30	3963	30.70
Political Science	55	9183	7239	78.83	1944	21.17
Sociology	36	5587	4217	75.48	1370	24.52
Total	433	58455	46502	79.55	11953	20.45

*** JMC- Journalism and Mass Communication ** LIS- Library and Information Science**

This study investigates how print and web citations are used in Indian theses. A total of 58455 citations were provided for the 433 selected theses (Table 6.6). Print citations comprise the majority of the 58455 citations (79.55%), with web citations coming in second place at (20.45%). The greatest web citations of theses have come from Journalism and Mass Communication (30.91%), with Library and Information Science (30.70%) and Sociology (24.52%) trailing closely behind. The table shows that web citations are not more common among researchers in Social Science.

Table: 6.7. Distribution of Web citations by Year (cross-tabulated by subject)

	Number of web citations				
Subject	**2020**	**2021**	**2022**	**2023**	**Total**
Economics	6	593	1199	1903	3701
*JMC	2	200	501	272	975
**LIS	5	225	2101	1632	3963
Political Science	3	225	783	933	1944
Sociology	0	143	661	566	1370
Total	16	1386	5245	5306	11953

*** JMC- Journalism and Mass Communication * *LIS- Library and Information science**

Table 6.7 presents the distribution of web citations in various subjects cross-tabulated by year. It is observed that the highest number of web citations was used in Library and Information Sciences (2101) in the year 2022 and the lowest (02) in Journalism and Mass Communication in the year 2020. For the entire period of study, theses in Library and Information Sciences had the highest number of web citations (3963), and theses in Journalism and Mass Communication had the lowest (975).

There were no web citations in Sociology theses in the year 2020. There is a lot of variation in the distribution of web citations among the subjects throughout the study in different years.

Table: 6.8. Distribution of Web citations by top-level domains (cross-tabulated by subjects)

Domains	**Economics**	***JMC**	****LIS**	**Political Science**	**Sociology**	**Total**	**%**
com/co	740	345	533	793	220	2631	22.01
org	677	179	511	500	204	2071	17.33
edu	114	45	527	33	49	768	6.43
gov	311	30	113	156	81	691	5.78
Country code Domains (.ca, .ir, .au, .uk, .nz etc.)	258	109	105	108	67	647	5.41
net	135	50	146	45	49	425	3.56
ac	93	39	185	23	31	371	3.10
nic	116	9	19	72	13	229	1.92
int	27	11	11	4	27	80	0.67
info	2	3	9	9	2	25	0.21
ernet	0	1	1	0	0	2	0.02
DOI	1230	154	1803	201	625	4013	33.57
Total	3703	975	3963	1944	1368	11953	100.00

*** JMC- Journalism and Mass Communication ** LIS- Library and Information Science**

The Political Sciences theses have the highest number of web citations from the.com domain (793), indicating the availability of much of the scholarly content in commercial databases and websites. This is followed by Economics (740) in which a total of 22.01% of the web citations belong to the .com domain. This clearly shows a substantial volume of scholarly literature in Political science (793) has the highest number of web citations under the .com domain, closely followed by theses in Economics (740). The lowest number of web citations are found under the *ernet* domain.

In the present study, 33.57% of DOIs are found in theses. This clearly shows that web resources in the Social Sciences have more DOIs. The highest number of DOIs is found in Library and Information Science (1803), while the lowest number is found in Journalism and Mass Communication (154).

The data displayed in Table 6.9 indicates that most cited URL citations were from HTML/htm files, accounting for 83.80% of the total, with PDF files coming in second place, representing 10.83% of the total. The data indicates that over 90% of the URL citations are from HTML and PDF files. The remaining 11 file formats identified in this study cumulatively accounted for 5.38% of the total.

Table: 6.9. Distribution of Web citations by file format (cross-tabulated by subject)

File Formats	Economics	*JMC	**LIS	Political Science	Sociology	Total	Percentage
html/htm	2946	855	3427	1633	1156	10017	83.80
pdf	538	77	368	135	176	1294	10.83
doc	2	1	2	3	0	8	0.07
php	36	12	88	20	16	172	1.44
cfm	1	0	0	0	0	1	0.01
nsf	4	4	8	1	1	18	0.15
asp/x	91	0	29	27	11	158	1.32
ppt	0	0	1	0	0	1	0.01
txt	0	0	1	1	0	2	0.02
shtml	2	0	3	3	0	8	0.07
cgi	6	1	32	1	4	44	0.37
cms	41	12	2	42	1	98	0.82
ece	36	13	2	78	3	132	1.10
Total	3703	975	3963	1944	1368	11953	100.00

*** JMC- Journalism and Mass Communication * *LIS- Library and Information science**

Table 6.10. Distribution of Web citations by path depth (cross-tabulated by subject)

Path depth	Economics	*JMC	**LIS	Political Science	Sociology	Total	Percentage
PD = 0	80	17	34	6	32	169	1.41
PD = 1	729	397	376	344	206	2052	17.17
PD = 2	1808	261	2631	622	767	6089	50.94
PD = 3	454	124	392	336	170	1476	12.35
PD = 4	256	80	189	372	94	991	8.29
PD = 5	251	59	196	178	59	743	6.22
PD = 6	74	23	75	54	21	247	2.07
PD = 7	27	10	41	18	13	109	0.91
PD = 8	13	1	12	8	5	39	0.33
PD = 9	8	2	7	4	1	22	0.18
PD>10	3	1	10	2	0	16	0.13
Total	3703	975	3963	1944	1368	11953	100.00

*** JMC- Journalism and Mass Communication ** LIS- Library and Information Science**

In the present study, path depth refers to the depth of pages inside the website or the number of hierarchical levels a URL consists of. For instance, the URL www.tumkuruniversity.ac.in has a path depth of 0, whereas the URL www.tumkuruniversity.ac.in/index.php has a path depth of 1. In this context, the study attempted to determine the path depth of the web citations. The data is presented in Table 6.10.

Table 6.10 illustrates the distribution of web citations based on path depth. Out of 11953 web citations, URLs with path depth PD=2 (50.94%) were frequently cited, followed by URLs with depth of PD=1 (17.17%)) and path depth PD=3 (12.35%). The least number of cited URLs have a path depth of PD>10, which accounts for only (0.13%).

7. Discussion and Conclusion

According to the study, throughout the last four years (2020–23), most theses uploaded to Shodhganga are from Economics (51.50%), while the least amount has come from JMC (3.70%). Nonetheless, from 2020 to 2023, there was a notable increase in the number of theses uploaded to Shodhganga. Among 58455 references found in 433 theses, 20.45% are web citations that come in second to print references (79.55%). Out of all the subjects, JMC theses had the highest number of web citations (30.91%), followed by sociology (24.52%) and library and information science (30.70%). The study shows that Social Science research academics do not find web citations more common. Additionally, a decrease in references is observed in 2023 compared to 2020, when the total number of references utilized in various theses appears to have increased.

A total of 433 theses mentioned 58455 references. Out of the entire 58455 references, print references make up the majority (79.55%), followed by web citations of the 2631 (22.01%) web citations; the most popular domain, com/.co, is well-represented throughout all topic areas. Furthermore, the HTML/htm file type accounted for the most significant percentage of websites (83.80%), followed by the pdf and php file formats (10.83%) and 1.44%, respectively. With 6089 (50.94%) references, a path depth of 2 has the most web references among all subjects.

From the discussion above, it can be noticed that Indian Social science research academics have only cited 20.45% of web citations. This suggests that the websites might not be reliable sources of information for them. There are several issues with web citations, including that some websites are just taken down when their domain names expire. The webmaster may also choose to delete the pages from the Internet in other situations. New file structures are used to construct certain websites, and certain websites might relocate to different web addresses. As a result, authors should verify the availability of web sources and include complete bibliographic information in their citations to web content. Before mentioning Web sources in academic writings, authors should ensure they are still accessible.

Acknowledgement

The research work was supported by the Indian Council of Social Science Research (ICSSR), New Delhi, INDIA.

References

[1] Bansal, S. (2021). Decay of URL references cited in DESIDOC Journal of Library & Information Technology. *DigitalCommons@University of Nebraska - Lincoln.* Retrieved from https://digitalcommons.unl.edu/libphilprac/5811

[2] Bapte, V. D. (2017). DESIDOC Journal of Library and Information Technology (DJLIT): A bibliometric analysis of cited references. *DESIDOC Journal of Library & Information Technology, 37*(4), 264. https://doi.org/10.14429/djlit.37.4.10712

[3] Carnevale, R. J., Aronsky, D. (2007). The life and death of URLs in five biomedical informatics journals. *International Journal of Medical Informatics, 76*(4), 269–273.

[4] Jan, R. (2009). Citation analysis of *Library Trends. Directory of Open Access Journals (DOAJ).* Retrieved from https://doaj.org/article/03e265d3bd124210b7ef16e0073d44a4

[5] Kaur, A., Rattan, G. K. (2018). Citation analysis of theses in economics submitted to Punjabi University, Patiala during 2000–2014. *DESIDOC Journal of Library and Information Technology, 38*(3), 192. https://doi.org/10.14429/djlit.38.3.12474

[6] Kumar, B. T. S., Prithviraj, K. (2014). Bringing life to dead: Role of Wayback Machine in retrieving vanished URLs. *Journal of Information Science, 41*(1), 71–81. https://doi.org/10.1177/0165551514552752

[7] Kumar, D. V., Kumar, B. T. S. (2017). Prevalence of URLs in Library and Information Science (LIS) literature: A citation analysis. *Collnet Journal of Scientometrics and Information Management, 11*(2), 287–297. https://doi.org/10.1080/

09737766.2017.1299390

[8] Kumar, R. A. (2017). Web citations in Indian theses: Exploring the usage and attributes. *International Journal of Library and Information Studies, 7*(1), 245–249.

[9] Kumar, D. V., Kumar, B. T. S. (2017). Recovery of vanished URLs: Comparing the efficiency of Internet Archive and Google. *Malaysian Journal of Library & Information Science, 22*(2), 31–43. https://doi.org/10.22452/mjlis.vol22no2.3

[10] Mahajan, P., Kumar, A. (2017). Citation analysis of doctoral theses in the field of sociology submitted to Panjab University, Chandigarh (India) during 2002–2012. *Library Philosophy and Practice (E-journal).* Retrieved from http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=4634&context=libphilprac

[11] Manjunatha, G., Kumar, B. T. S., Lakshmana, H. (2020). Longevity of URL citations cited in LIS journal articles: A webometric study. *Library Philosophy and Practice (E-journal).* Retrieved from https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=7336&context=libphilprac

[12] Niveditha, B., Kumbar, M. (2020). Web citation analysis of Library and Information Science and Communication and Media Studies journals: A comparative study. *Collnet Journal of Scientometrics and Information Management, 14*(2), 335–348. https://doi.org/10.1080/09737766.2021.1915721

[13] Patel, A. K., Singh, M., Singh, K., Patel, A. K., Varma, A. K., Kuri, R. (2021). Visualizing publication trends in Webology Journal: A bibliometric review based on the Scopus database (2006–2020). *Library Philosophy and Practice (E-journal).* Retrieved from https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=11112&context=libphilprac

[14] Prithviraj, K., Kumar, B. T. S. (2015). Web citation trends in Indian LIS journals: A citation analysis. *Collnet Journal of Scientometrics and Information Management, 9*(2), 295–310. https://doi.org/10.1080/09737766.2015.1069968

[15] Sheoran, M., Kumar, D., Kumar, V., Verma, D. (2018). Understanding the trends of marketing research and its future directions: A citation analysis. *The Bottom Line, 31*(3/4), 191–207. https://doi.org/10.1108/bl-04-2018-0022

[16] Loan, F. A., Shah, U. Y. (2020). The decay and persistence of web references. *Digital Library Perspectives, 36*(2), 157–166. https://doi.org/10.1108/DLP-02-2020-0013

[17] Teixeira da Silva, J. A., Nazarovets, M. (2023). Archiving website-based references in academic papers: Problems caused by reference rot, potential solutions and limitations. Retrieved from https://www.researchgate.net/publication/372086239_Archiving_website-based_references_in_academic_papers_Problems_caused_by_reference_rot_potential_solutions_and_limitations.pdf