Annals of Library and Information Studies
Vol. 60, December 2013, pp. 284-295

URLs as references in Indian LIS conference papers: an Analysis

K.R. Prithviraj[a] and B.T. Sampath Kumar [b]

[a]Research and Teaching Assistant in Library & Information Science, Directorate of Distance Education, Kuvempu University, Jnana Sahyadri, Shivamogga, Karnataka, India, Email:prithviraj.kr@gmail.com

[b]Associate Professor, Department of Studies and Research in Library and Information Science, Tumkur University, Tumkur, Karnataka, India, Email:sampathbt_2001@rediffmail.com

Received 16 July 2013; revised and accepted 14 November 2013

The paper compares the characteristics of URLs cited in Indian LIS conference proceedings papers.A total of 15,745 references appended to 1,700 articles published in three Indian LIS conference proceedings published during 2001-2010 were selected. From these references we extracted a total of 5698 URLs and were further classified according to their top level domains, file formats and path depths for further analysis.The results showed that the percentage of articles with at least one URL increased from 39.10% in 2001 to 91.67% in 2010. There was a constant and continuous increase in the number of articles with URLs over the years during 2001-2010. Of the 1,700 articles published in conference proceedings, there were 1,011 (59.47%) articles with URLs. Study also reveals the fact that, of the 5,698 URLs, more than 50% were shared by .org and .com domains which accounted for 1,799 (31.57%) and 1,474 (25.87%) URLs respectively. The highest percentage of cited URLs belonged to HTML (68.50%) followed by .pdf files (8.86%). The path depth levels 0 (no path), 2 and 3 collectively accounted for 67.67% of the extracted URLs. URLs with path depth 1 and 4 put together accounted for 25.31% of all the 5,698 URLs.

Keywords: Internet; URLs; citations; Conference proceedings; India

Introduction

Since the quasi-miraculous emergence of the Web in 1990s, there has been a continuous increase in the volume of scholarly resources in electronic form, such as e-books, e-journals, e-databases, e-theses and dissertations, e-prints of research papers, and the like. These resources have provided a scope for researchers and authors in various subject fields and stimulated their research productivity. Library and information science (LIS) was no exception to this[1]. With the growth of scholarly publication on the web all over the world, the use of web links or URLs have become common in journal articles, conference papers and other scholarly publications in all disciplines. Thus the increased amount of information available on the web has influenced the corresponding use of URLs by authors in their scholarly works[2-7].The present study investigates the use of URLs as citations in the articles of Indian LIS conference proceedings published during 2001 to 2010. The study also tries to assess the trends in the URL citations and compares the characteristics of URLs cited in Indian LIS conference proceedings.

Related literature

Citations to web resources in scholarly publications have been studied since as early as the mid-1990s. Lawrence[8] analysed 1,19,924 conference articles in computer science and related disciplines and found that highly cited articles were substantially more likely to be freely available on the Web. The same study findings showed that the mean number of citations to offline articles was 2.74 while the mean number of citations to online articles was 7.03: i.e, 2.6 times greater than the number for offline articles. Lawrence et al.[2] investigated URLs contained in research articles from the Research Index database (CitSeer). They analysed 2,70,977 articles and extracted 67,577 URLs to examine their accessibility on the Web.

Casserly and Bird[9] conducted a study of LIS scholarly journal articles published between 1999 and 2000, and found that among the 35,689 references cited in 1,425 articles, a total of 3,582 of the references originated from the Web. Hester *et al.*[10] conducted citation analysis of 45 journals in Oncology, found that articles with web citations increased from 9% in 2001 to 16% in 2003. Sellitto[11] analyzed 123 articles on Web based education and training research obtained from the Aus Web series of conferences between 1995 and 2003. Of the 2,168 references obtained, almost 50% were URLs, giving the average number of web citations to be 8.5 per article.

Patil et al.[12] conducted a study on citation analysis of the articles published in Indian Library Association's Conference Proceedings of 2003. Authors analyzed 790 references cited in 94 research articles of which 196 (24%) were web citations, with web domain .com contributing the highest number 48 (24%) followed by .org 44 (22%). Another study by Bhat and Sampath Kumar[4] on citation analysis of research articles from scholarly electronic journals in the field of LIS showed that there was an increasing trend in the use of web citations. 81.05% of articles published in the selected nine journals have web citations. Out of 25,730 references 56.5% of references were print journal references and 43.5% were web references. Janakiramaiah and Doraswamy[13] analyzed 689 web citations selected from the 1,885 references cited in the conference articles published in the Convention on Automation of Libraries in Education and Research Institutions (CALIBER) of 2005 and 2006. In this study they found that 36.55% were web citations out of the total 1,885 references. The highest number of web citations (28.45%) were from organizational (org) websites. Riahinia et.al.[14] selected all volumes of six LIS journals ranked by ISI Thomson Reuters published during 2005-2008. From 1,181 papers, 37,791 citations were recorded. Of the 37,791 citations, 4,840 (12.8 percent) were web citations. The means per articles of web and print citations were 4.09, and 27.9, respectively.

In the same year Nagaraja et al.[15] collected a total of 1,133 articles published from 2005-2007 in PLoS Medicine (Public Library of Science Medicine). The 1,133 articles contained 28,177 references, with 2,503 (8.9%) identified as URLs. Non-research articles accounted for a substantially higher percentage of URL references (17.4%) compared to research articles (4.2%). Sampath Kumar and Prithviraj[7] analysed 673 web citations from 350 articles published in IATLIS conference proceedings during 2001 to 2008. The articles contained a total of 2,711 references, with 24.82% (673) of references citing to web resource. The average number of web citations per article increased substantially from 0.25 (in the year 2001) to 3.02 (in the year 2008). The web citations as a percentage of all citations also significantly increased during 2001 to 2008 (4.01 in 2001 to 27.25 in 2008). During 2001 to 2008, web citations in the scholarly articles increased by approximately six fold.

Saberi and Abedi[16] analysed 3,734 unique web citations in 748 articles published in five open access journals as indexed by the Institute for Scientific Information (ISI), from 2002 to the end of 2007. The overall average of web citations per article was 4.9. Sampath Kumar and Manoj Kumar[17] conducted a study on the availability and persistence of URL citations in two LIS open access scholarly journals. A total of 2,890 URL citations cited in 689 research articles published in LIS journals spanning a period of 14 years (1996-2009) were extracted. The research findings indicated that 57.61 percent (397 of 689) of articles have URL citations and percentage of URL citations increased from 5.98 percent in 1996 to 27.79 percent in 2009. Sampath Kumar and Vinay Kumar[18] conducted a study on the availability, persistence and half life of URL citations cited in two Indian LIS scholarly journals. Authors found that a total of 472 research articles published in two Indian LIS journals for the years 2002–2010. These research articles contained a total of 6820 references citations (mean = 14.44), with only 18.91% (1290) of references citing a URLs. The percentage of URLs by year varied from a low of 9.54 in the year 2007 to high of 30.82 in the year 2008. From the above discussion it is very clear that the use of URLs as citations is increasing in the scholarly literature. Table1 summaries percentage URLs used as citation in previous studies conducted by various authors.

Objectives of the study

The main purpose of the present study is to add to the body of knowledge about the changing landscape of scholarly communications by examining the citations to web sources (URLs) included in articles published in Indian Library and Information Science (LIS) conference proceedings during 2001 to 2010. The main research objectives of the study are:

- To compare the proportion of URLs used as citations in the articles of Indian LIS conference proceedings;
- To identify and compare relationship between the authorship patterns and URL citations;
- To know the top level domain and file format constitutes the highest number of URLs; and
- To measure the path depth and character length of URL citations.

Hypothesis

- URLs are most preferred citations cited in Indian conference proceedings.
- There is an association between the authorship pattern and use of URLs.

Table 1—Use of URLs as citations in earlier studies

Sl. no.	Author	Year	Discipline	Percentage of URLs
1.	Harter & Kim[19]	1996	Scholarly articles citations	1.92 %
2.	Davis & Cohen[20]	2001	Students term paper microeconomics	9% in 1996 22% in 2000
3.	Herring [21]	2002	Online journals of different disciplines	16%
4.	Casserly & Bird [9]	2003	Library and Information Science Journals	10%
5.	Hester et al.[10]	2004	Oncology journals	0.63%
6.	Sellitto[11]	2004	Conference proceedings	48.1%
7.	Kushkowski[22]	2005	Electronic theses	3.54%
8.	Olfson & Lawrence [23]	2005	Medical and science journals	0.7% in 2001 2.31% in 2004
9.	Maharana et al.[1]	2006	Proceedings of SIS-2005	34.88%
10.	Megnigbeto[24]	2006	Library and Information Science Students dissertation	4.85%
11.	Patil et al. [12]	2007	LIS Conference proceedings	24.81%
12.	Ducut et al.[25]	2008	MEDLINE record	10,208
13.	VeenaBhat & Sampath Kumar [4]	2008	E-journals in LIS	43.52%
14.	Wu [5]	2009	LIS Journal and Computer Sci. Journal	8.41%
15.	Chen et al.[26]	2009	Journals of different disciplines	0.24%
16.	Yang et al.[27]	2010	Journals Chinese social science citation index	3.16%
17.	Janakiramaiah & Doraswamy[13]	2011	Conference proceedings	36.55%
18.	Riahinia et al.[14]	2011	Library and Information Science Journals	12.8%
19.	Nagaraja et al. [15]	2011	PloS Medicine	8.9%
20.	Saberi et al. [28]	2011	Journal of Artificial Societies and Social Simulation	7.51%
21.	Sampath Kumar & Prithviraj [7]	2012	LIS Conference proceedings	24.82%
22.	Mardani, [29]	2012	Chemistry articles by Iranian researchers indexed in SCI	24.90%
23.	Sadat-Moosavi et al.[30]	2012	Library and Information Science Journals	16.28%
24.	Sampath Kumar &Manoj Kumar [17]	2012a	Library and Information Science Journals	18.77%
25.	Sampath Kumar &Manoj Kumar [31]	2012b	Two science and social science journals	2.98 %
26.	Sampath Kumar & Vinay Kumar [19]	2013	Library and Information Science Journals	18.91%
27.	Jalalifard, M et al.[32]	2013	Seven medical journals	0.89%

- The bibliographical description of URLs (content of URLs) varies among the conference proceedings.
- The top level domain of URLs varies among the conference proceedings.

Methodology

The data for the present study were drawn from three Indian LIS conference proceedings published in printed form. Every year in India there are at least half a dozen national associations/institutions which conduct annual conferences in the field of LIS. In this study we considered the proceedings of (i) CALIBER conference conducted by INFLIBNET, (ii) NACLIN conference conducted by DELNET and (iii) ILA conference conducted by Indian Library Association published during 2001-20010. These three conference proceedings were considered for two reasons. Firstly these conference proceedings have a wide coverage of participants in the field of Library and Information Science. Secondly these conference proceedings have longer period of history.

All research articles published during the 10-year period (2001-2010) were extracted and saved. Editorial notes or book reviews were excluded. Totally 15,745 references were selected from 1,700 articles published in the selected Indian LIS conference proceedings. Only references that appeared as a list at the end of the article under the bibliography or hypertext reference section were considered. Expanded bibliographies, end notes, foot notes, e-mail links and annotations were not considered and were not tested or counted in our dataset. After selecting all references appended to the articles published in conference proceedings, URLs were extracted for further analysis.

Analysis

Distribution of citations into types of information sources

A total of 1,700 articles were published in LIS conference proceedings during 2001-2010. A close look at the data clearly indicates that the highest number of articles (220) were published in 2007 and

the least number of articles (156) were published in 2001. CALIBER and ILA conference were not held in the year 2010 and hence only articles published in NACLIN were considered for analysis. Thus the number of articles (24) published in the year 2010, were comparatively very less.

The data presented in Table 2 shows that there is an upward trend in the growth of scholarly literature in LIS field. It was also found that there is a steep growth in the number of conference articles in the second half (2006 to 2009). A total of 15,745 citations were cited in 1,700 articles published during 2001-2010, with an average of 9.26 citations per article. The average citation per article was also at an increasing trend ranging from 7.47% in 2001 to 12.54% in 2010. This analysis clearly indicates that the LIS scholarly articles were at constant increase in the later years of the decade studied. This trend also automatically increased the number of citations in the scholarly writings.

A bird's eye view of the distribution of the 15,745 citations into different types of sources cited in the 1,700 articles, showed that highest number of citations, i.e. 5,698 (36.19%) were URLs. It was followed by citations from other information sources, such as Journals 3,647 (23.16%), Books 3,322 (21.10%), and Conference Proceedings 1,116 (7.09%). However, only 61 (0.39%) and 49 (0.31%) citations had come from Theses and Dissertations respectively. And the remaining 888 (5.64%) citations were from 'others' category. It is evident from the fact that 36.19% of all the cited references in this study were URLs which occupied the first place followed by journals and books. The correlation analysis also indicates that the number of citations and URLs are positively correlated and the correlation is statistically significant (r=.927, p=.000).

Distribution of articles, citations and URLs

The data presented in Table3 clearly shows that the percentage of articles with at least one URL increased from 39.10% in 2001 to 91.67% in 2010. There was a constant and continuous increase in the number of articles with URLs over the years during 2001-2010. Of the 1,700 articles published in conference proceedings, there were 1,011 (59.47%) articles with URLs.

The average number of URLs per article ranged from a low of 1.99 in 2001 to a high of 5.63 in 2010. The average number of URLs per article was 3.35 across all the 1,700 conference articles. The data regarding the average print citations and URLs in conference proceedings is also represented in the form of graph (Fig. 1). Further, the year 2006 and 2010 witnessed the highest percent of URLs with 47.24% and 44.85% respectively, which were comparatively higher than the decadal percentage of URLs (36.19%). The percentage of URLs had increased from the year 2001 (26.70%) to 2010 (44.85%). The correlation analysis indicates that the there is a negative co-relation between the age and percentage of URLs (R=-769, p=.015) and the correlation is also statistically significant. This indicates that the conference proceedings published in recent years contained more number of URLs compared older ones.

Table 2—Distribution of citations in to types of information sources

Publication year of the conference proceedings	Total no. of articles	Total no. of citation	Average citations per article	Distribution of Citations into types of information sources								
				Journals	Books	Reports	Workshops /Seminars	Conference proceedings	URLs	Theses	Dissertation	Others*
2001	156	1165	7.47	349	277	24	58	62	311	1	9	74
2002	174	1453	8.35	365	339	31	50	148	387	7	2	124
2003	182	1474	8.10	333	337	47	69	97	449	11	2	129
2004	171	1519	8.88	289	337	33	84	120	537	7	3	109
2005	168	1636	9.74	343	412	24	58	150	558	8	4	79
2006	195	1867	9.57	370	268	22	100	133	882	4	5	83
2007	220	2022	9.19	442	435	52	78	132	775	2	9	97
2008	214	2179	10.18	524	451	46	82	130	804	14	7	121
2009	196	2129	10.86	558	423	26	57	127	860	7	8	63
2010	24	301	12.54	74	43	2	21	17	135	0	0	9
All years	1700	15745	9.26	3647	3322	307	657	1116	5698	61	49	888
Percentage of citations (N=15745)				23.16	21.10	1.95	4.17	7.09	36.19	0.39	0.31	5.64

*Others: archives, syllabus, broachers, news papers, etc,

Use URLs as citations in conference proceedings

A bird's eye view of the data presented in Table4 clearly shows there were 59.74% of articles with URLs. The percentage of articles with URL found highest in CALIBER proceedings (72.28%) followed by NACLIN (64.13%) and ILA (45.76%). The average number of URLs per article was 3.35 across all the 1,700 conference articles. The data regarding the

Table 3—Distribution of articles, citations and URLs

Publication year of the conference proceedings	Total no. of articles	No. of articles with URLs	Percentage of articles with URLs	Total no. of citations	Total no. of print citations	No. of print citations per article	Total no. of URLs	No. of URLs per article	Percentage of URLs
2001	156	61	39.10	1165	854	5.47	311	1.99	26.70
2002	174	82	47.13	1453	1066	6.13	387	2.22	26.63
2003	182	71	39.01	1474	1025	5.63	449	2.47	30.46
2004	171	97	56.73	1519	982	5.74	537	3.14	35.35
2005	168	113	67.26	1636	1078	6.42	558	3.32	34.11
2006	195	130	66.67	1867	985	5.05	882	4.52	47.24
2007	220	138	62.73	2022	1247	5.67	775	3.52	38.33
2008	214	153	71.50	2179	1375	6.43	804	3.76	36.90
2009	196	144	73.47	2129	1269	6.47	860	4.39	40.39
2010	24	22	91.67	301	166	6.92	135	5.63	44.85
All years	1700	1011	59.47	15745	10047	5.91	5698	3.35	36.19
Average citations per article				9.26					
Percentage of citations of all the years					63.81		36.19		

Table 4—Use URLs as citations in conference proceedings

Publication year of the conference proceedings	Total no. of articles	No. of articles with URLs	Percentage of articles with URLs	Total no. of citations	Total no. of print citations	No. of print citations per article	Total no. of URLs	No. of URLs per article	% URLs
CALIBER	624	451	72.28	6144	3655	8.10	2489	3.99	40.51
ILA	708	324	45.76	5949	4315	13.32	1634	2.31	27.47
NACLIN	368	236	64.13	3652	2077	8.80	1575	4.28	43.13
Total	1700	1011	59.47	15745	10047	9.94	5698	3.35	36.19

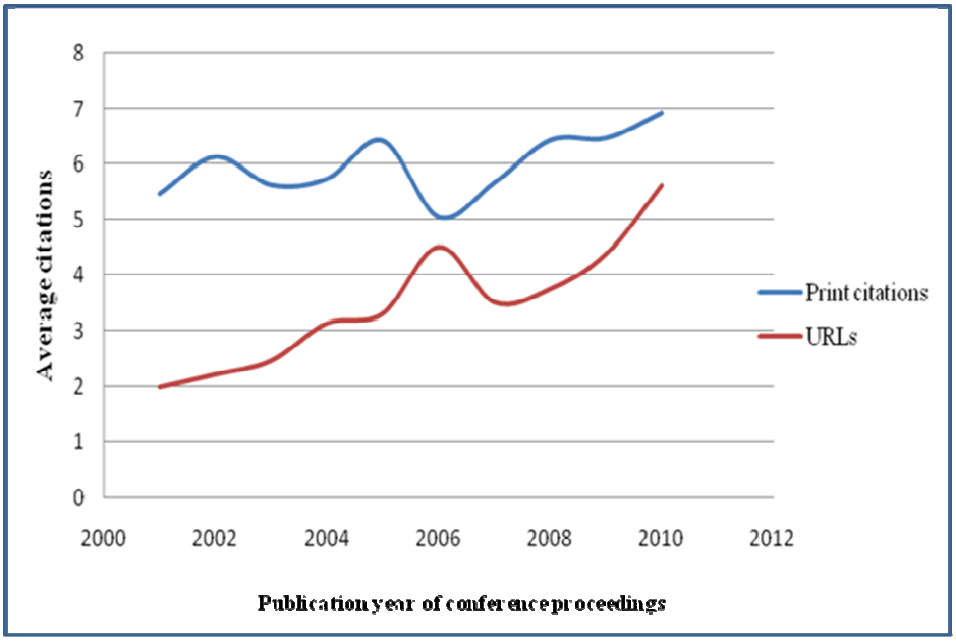

Fig. 1—Average print citations and URLs

Percentage of URLs in conference proceedings is also represented in the form of graph (Fig. 2).

Authorship pattern and citations

Table 5 illustrates the authorship pattern of the articles published in the selected LIS conference proceedings and as well as their citation pattern. It reveals that the articles with single author were more in number (41.71%) followed by the articles with two authors (39.41%). The third, fourth and fifth place goes to the articles with three, four and more than five authors, which accounted for 15.35%, 3.18% and 0.35% of the total articles respectively. There is a definite relation between the authorship pattern and the citation pattern. The further analysis of the citation data in Table 4 clearly shows that the number of citations per article goes on increasing with the increase in the number of authors per article. The correlation analysis also indicates that the number of authors and the average URLs per article is positively correlated (r=.544) but the correlation is not statistically significant (p=.343).

The conference articles with five and more authors had the highest average number of print citations (8.17) and as well as URLs (4.17) per article, it is also presented in Fig.3. The articles with single authorship had the least average number of print citations (5.74), but when it comes to the number of URLs, the articles with three authors had the least average number of

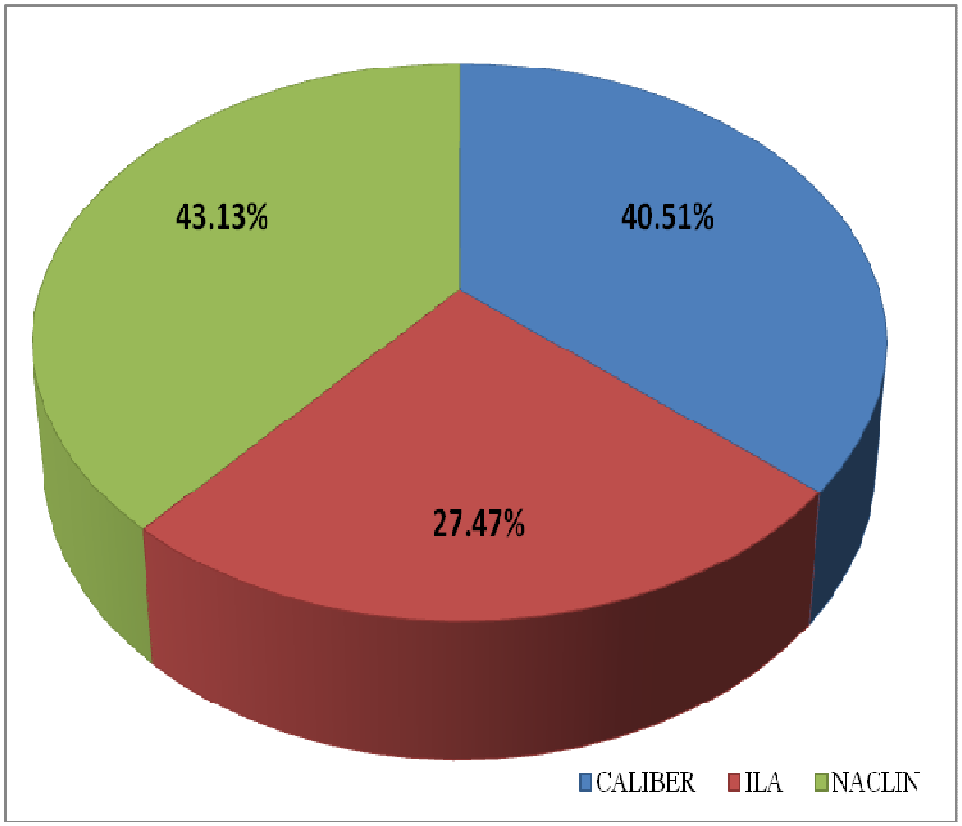

Fig. 2—Percentage of URLs in conference proceedings

Table 5—Authorship pattern and citations

No. of authors	Total no. of articles	Percentage	Print citations		URLs	
			Number	Average per article	Number	Average per article
Single	709	41.71	4071	5.74	2450	3.46
Two	670	39.41	4071	6.08	2274	3.39
Three	261	15.35	1528	5.85	759	2.91
Four	54	3.18	328	6.07	190	3.52
Five and above	6	0.35	49	8.17	25	4.17
Total	1700	100	10047	5.91	5698	3.35

URLs (2.91). We also tested the association between the authorship pattern and the use of print citations as well as URLs using Chi-square analysis. We found the there is a significant association between the authorship pattern and citation behaviors of authors (X^2=14.890, DF=4, p=.005, c=.031).

Contents of URLs

A citation is considered complete if it includes maximum information about the web document i.e name of the author, title, publication, publisher, date of publication, URL, date of access etc., The web citation content noted in the references list varies from a 'URL only' to a 'URL accompanied by partial or complete bibliographic information'. A total of 5,698 URLs found in the selected conference proceedings were classified on the basis of the web citation content and the resultant data is presented in Table6 for further analysis.

The analysis of data presented in Table 6 shows that 33.15% of the 5,698 URLs, contained only URLs, whereas 56.83% contained URLs with partial bibliographic information and only 10.02% of the URLs contained URLs with complete bibliographic information (Fig.4). URLs cited in NACLIN conference proceedings had more number of web sources with only URL address (42.60%) followed by ILA (33.35%) and CALIBER (27.04%). While only 8.51% of web sources cited in NACLIN had full bibliographical details. This clearly indicates that the content of the URLs used as citation varied among conference proceedings (X^2=1.086, DF=4, p=.000, c=.137).

Domains associated with URLs

In this study, seven different types of top-level domains have been identified among the URLs. They were .org, .com, .edu, .ac, .gov, .net and .ernet and all the other domains that accounted for less than 1% each of all the URLs of the URLs studied and those which do not fall into any of these seven major categories were grouped into the "others" category.

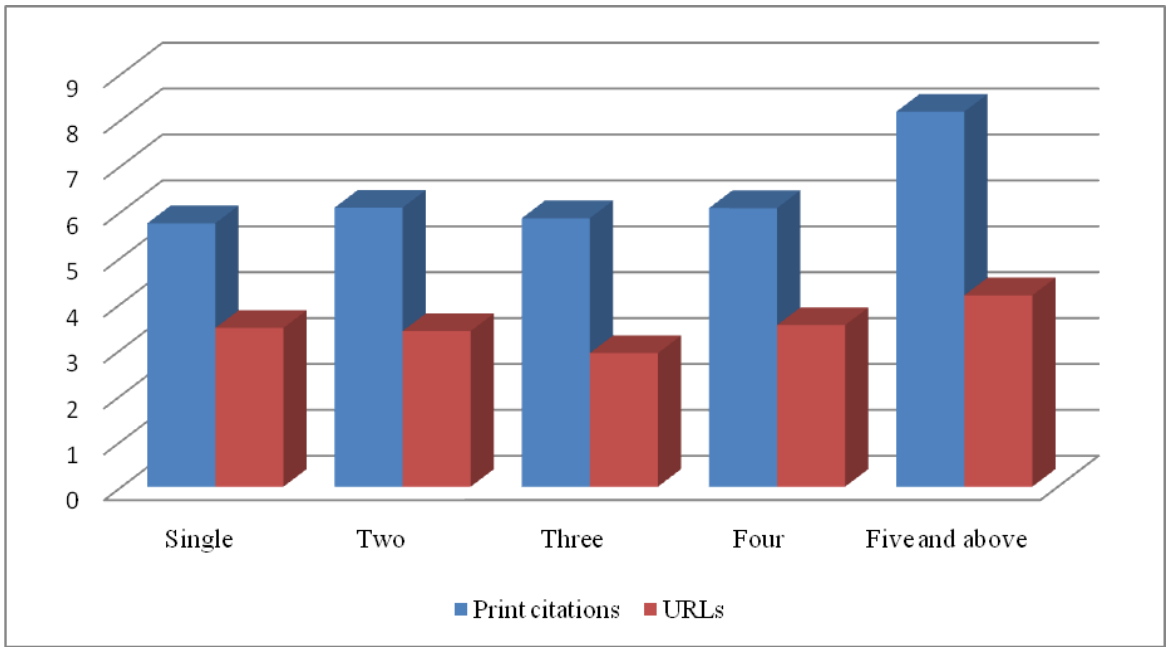

Fig. 3 —Authorship pattern and average citations

Table 6—Contents of URLs

Conference proceedings	Total no. of URLs	Only URL	%	URL & partial bibliographic information	%	URL & complete bibliographic information	%
CALIBER	2489	673	27.04	1559	62.64	257	10.33
ILA	1634	545	33.35	909	55.63	180	11.02
NACLIN	1575	671	42.60	770	48.89	134	8.51
Total	5698	1889	33.15	3238	56.83	571	10.02

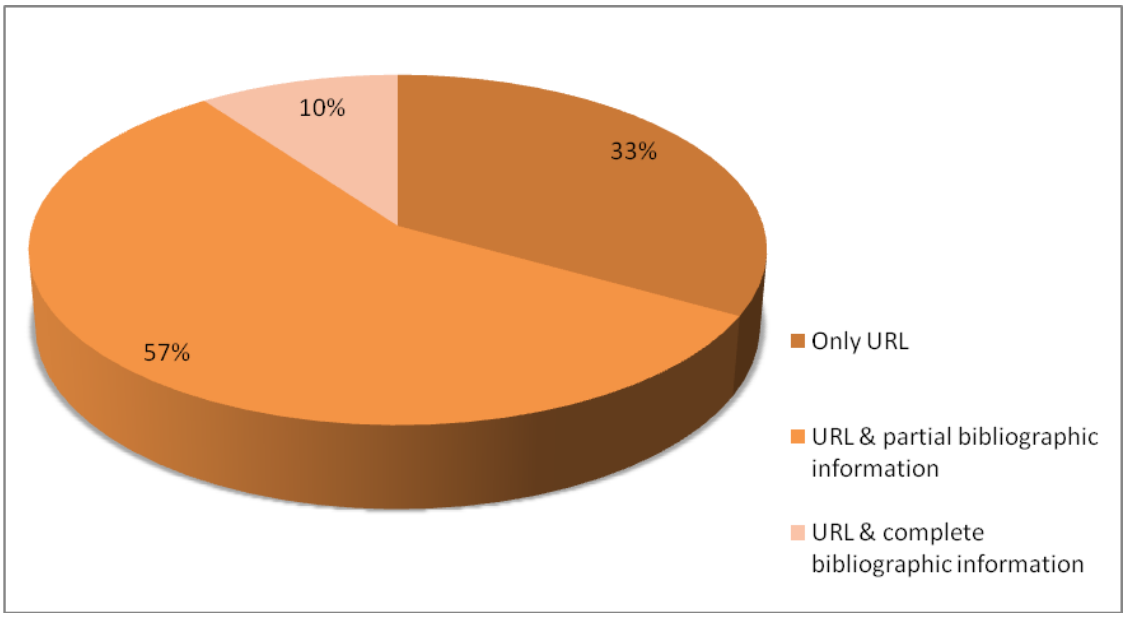

Fig. 4 —Contents of URLs

Table 7—Domains associates with URLs

Domains	CALIBER		ILA		NACLIN		Total	
	Number	%	Number	%	Number	%	Number	%
.org	784	31.50	516	31.58	499	31.68	1799	31.57
.com	757	30.41	393	24.05	324	20.57	1474	25.87
.edu	271	10.89	206	12.61	240	15.24	717	12.58
.ac	245	9.84	164	10.04	151	9.59	560	9.83
.gov	89	3.58	97	5.94	93	5.90	279	4.90
.net	70	2.81	39	2.39	51	3.24	160	2.81
.ernet	25	1.00	21	1.29	18	1.14	64	1.12
Other	248	9.96	198	12.12	199	12.63	645	11.32
Total	2489	100	1634	100	1575	100	5698	100

Each one of the entire 5,698 URLs identified in this study, from the selected LIS conference proceedings, have been classified and grouped into one or the other of these eight domain types. A close view of the data in Table7 reveals the fact that, of the 5,698 URLs, more than 50% were shared by .org and .com domains which accounted for 1,799 (31.57%) and 1,474 (25.87%) respectively. The other significant domains were .edu (12.58%), .ac (9.83%), .gov (4.90%), .net (2.81%), .ernet (1.12%). So thereby the .org, .edu and .com emerged as the 'big three' of the TLDs among the URLs of the total 5,698 URLs. This reveals that the data sources of more than 50% of the URLs in the present study were from websites of organizations and commercial domains. In support of our study the Chi-square analysis also indicates that there is an association between the conference domains associated with ULRs and conference proceedings (X^2=79.615, DF=14, p=.000, c=117).

File formats associated with URL's

The data presented in Table 8 shows that the highest percentage of cited URLs belonged to HTML files. Out of 5,698 URLs, 3,903 (68.50%) were of HTML files followed by .pdf files with 505 URLs (8.86%) at the second place and the files ending with number, were 232 (4.07%) which occupied the third place. The remaining file types had negligible number of URLs.

Path depth of URLs

In this study each of the 5,698 URLs were verified for their path depth and were classified and grouped in to the appropriate path depth levels from path depth of 0, 1, 2, 3, 4, 5 and any URL which has path depth level 6 and above were grouped into 6th path depth level. The analysis of data in the Table9 reveals that the path depth level 0 (no path), 2 and 3 collectively accounted for 67.67% of the extracted URLs. On the

Table 8—File formats associated with URL's

File formats	CALIBER		ILA		NACLIN		Total	
	Number	%	Number	%	Number	%	Number	%
.html/.htm	1727	69.39	1102	67.44	1074	68.19	3903	68.50
.pdf	175	7.03	158	9.67	172	10.92	505	8.86
.asp	38	1.53	21	1.29	17	1.08	76	1.33
.shtml	25	1.00	16	0.98	21	1.33	62	1.09
.php	21	0.84	12	0.73	12	0.76	45	0.79
.doc	16	0.64	9	0.55	11	0.70	36	0.63
.ppt	13	0.52	3	0.18	11	0.70	27	0.47
.txt	7	0.28	3	0.18	3	0.19	13	0.23
Ends with number	105	4.22	57	3.49	70	4.44	232	4.07
Others	362	14.55	253	15.49	184	11.69	799	14.02
Total	2489	100	1634	100	1575	100	5698	100

Table 9—Path depth of URLs

Path depth (PD)	CALIBER		ILA		NACLIN		Total	
	Number	%	Number	%	Number	%	Number	%
PD=0	643	25.83	405	24.79	367	23.30	1415	24.83
PD=1	305	12.25	231	14.14	225	14.29	761	13.36
PD=2	611	24.55	393	24.05	364	23.11	1368	24.01
PD=3	453	18.20	306	18.73	314	19.94	1073	18.83
PD=4	299	12.01	189	11.57	193	12.25	681	11.95
PD=5	95	3.82	61	3.73	70	4.44	226	3.97
PD≥6	83	3.34	49	2.99	42	2.67	174	3.05
Total	2489	100	1634	100	1575	100	5698	100

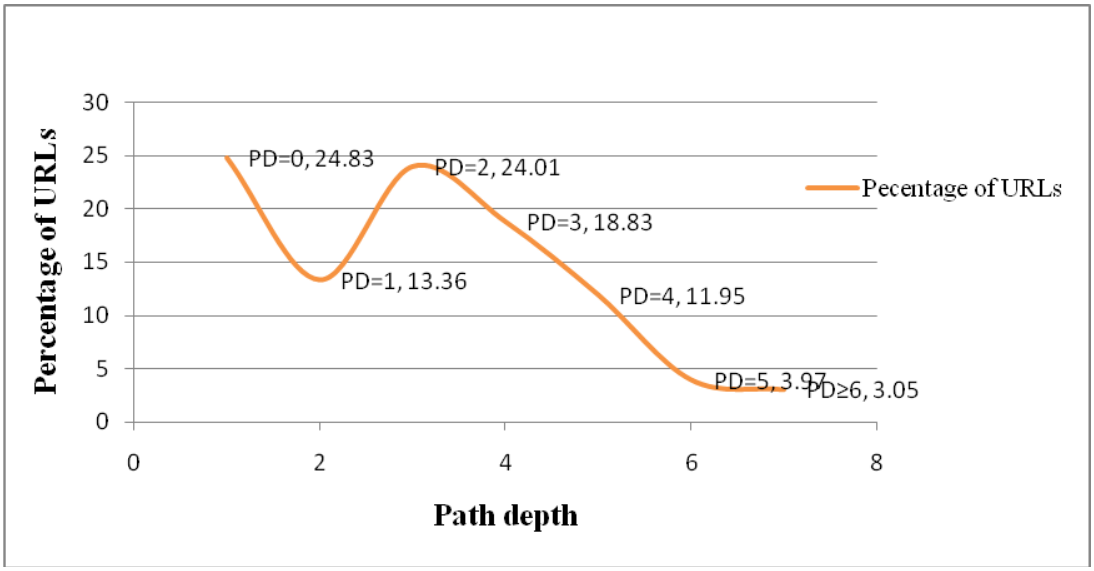

Fig. 5 —Path depth v/s percentage of URLs

other hand, URL with path depth 1 and 4 put together accounted for 25.31% of all the 5,698 URLs. The path depth 5, 6 and above accounted for 3.97% and 3.05% respectively (Fig.5). Thus, the URLs of the cited web resources in the LIS conference proceedings under the study were found densely concentrated at the URL path depth of1 0, 2 and 3, and fairly concentrated at path depth of 1 and 4 whereas, it was negligible at path depth 5, 6 and above.

URL characters length associates with URLs

The data presented in Table10 shows the distribution of 5,698 URLs into various character length groups. Analysis of the data reveals that the

Table 10—URL characters length associates with URLs

URL characters	CALIBER		ILA		NACLIN		Total	
	Number	%	Number	%	Number	%	Number	%
<20	233	9.36	131	8.02	124	7.87	488	8.56
21-30	467	18.76	306	18.73	289	18.35	1062	18.64
31-40	387	15.55	278	17.01	272	17.27	937	16.44
41-50	563	22.62	387	23.68	357	22.67	1307	22.94
51-60	410	16.47	259	15.85	267	16.95	936	16.43
61-70	196	7.87	142	8.69	122	7.75	460	8.07
71-80	122	4.90	71	4.35	65	4.13	258	4.53
>80	111	4.47	60	3.67	79	5.01	250	4.39
Total	2489	100	1634	100	1575	100	5698	100

Table 11—Testing of hypotheses

Sl. no.	Hypotheses	Test applied	P value and result
H1	URLs are most preferred citations cited in Indian conference proceedings.	Pearson's Correlation	P=.000 Hypothesis Accepted
H2	There is an association between the authorship pattern and use of URLs.	Chi-Square	P=.005 Hypothesis Accepted
H3	The bibliographical description of URLs (content of URLs) varies among the conference proceedings.	Chi-Square	P=.000 Hypothesis Accepted
H4	The top level domain of URLs varies among the conference proceedings.	Chi-Square	P=.000 Hypothesis Accepted

URLs with number of characters ranging from 41-50 were highest in number with 22.94% of URLs, followed by a group with 21-30 range with 18.64% URLs and 16.44% URLs in favor of the group with 31-40 characters. The least number of URLs were in the group with 80 and above characters, which accounted for only 4.39% of all the URLs. This clearly indicates that majority of URLs contained 41-50 characters length.

Testing of hypotheses

The present study formulated four hypotheses and these hypotheses were tested with the findings of the preset study. Pearson's correlation and Chi-Square test were applied to know the association between the variables and the result are shown in the Table11.

Conclusion

The present study examined the citation behavior of Indian LIS scholars by analyzing the citations appended to Indian LIS conference proceedings articles. Results of the study clearly showed that the percentage of articles with at least one URL increased from 39.10% in 2001 to 91.67% in 2010. Of the 1,700 articles published in conference proceedings, there were 1,011 (59.47%) articles with URLs. The average number of URLs ranged from a low of 1.99 in 2001 to a high of 5.63 in 2010. This indicates that use of URLs as citations in Indian LIS scholarly publications have been steadily increasing during last decade.

Use of URLs in selected conference proceedings showed that CALIBER proceeding (72.28%) had highest and ILA had least percent of articles with URLs (45.76%). The authorship pattern and use of print citations as well as URLs were examined and we found that the number of citations per article goes on increasing with the increase in the number of authors per article. The conference articles with five and more authors had the highest average number of print citations (8.17) and as well as URLs (4.17) per article. The articles with single authorship had the least average number of print citations (5.74) per article, but when it comes to the number of URLs the articles with three authors had the least average number of URLs (2.91) per article. Even though there is a positive correlation between the authorship pattern and the average number of URLs cited in article, the relation is not statistically significant (r=.544, p=.343).

When we analyzed the contents of the each URL it was found that 33.15% of the 5,698 web sources contained only URLs, whereas 56.83% contained URLs with partial bibliographic information and only 10.02% of them contained URLs with complete bibliographic information. Study recommends that author should provide the detailed bibliographical details of the web sources while citing them in the article. This will help the readers to retrieve the article when it moved and relocated to another URL address using various web archives or search engines.

Regarding the domains of the URLs, the study found that .org and .com domains were accounted for 1,799 (31.57%) and 1,474 (25.87%) respectively. Study also found that the highest percentage of cited URLs belonged to HTML files. Out of 5,698 URLs, 3,903 (68.50%) were of HTML files followed by .pdf files with 505 URLs (8.86%) at the second place. This indicates that authors preferred to cite URLs with HTML file extension. Another reason for citing more number of HTML files was that majority of information available on the web available in HTML format only.

Study showed that the percentage of URLs having path depth 0 (24.83%) and path depth 3 (24.01%) were more compared to path depth 5 and 6 (3.97% and 3.05% respectively). Earlier studies[22-26] have proved that the URLs with less path depth were likely to be more persistent than lengthy URLs. Thus it is recommended that the creator or author of a web page should use URL with shorter path depths.

The present study has some limitations which could be addressed in future research. In this study we considered only three Indian LIS conference proceedings published during 2001-2010. A total of 15,745 citations appended to 1,700 articles published in these conference proceedings articles were considered. A total of 5,698 URLs were extracted from reference and analyzed. However the data is not representative of the entire population of Indian LIS literature. Future researcher may extend this work by considering more number of conference proceedings published from India. Further, researcher may also focus on URL decay in LIS publications and recovery of decayed URLs using various web archiving tools.

References

1 Maharana B, Nayak K and Sahu N K, Scholarly use of web resources in LIS research: A citation analysis, *Library Review*, 55(9) (2006) 598-607.

2 Lawrence S, Coetzee F, Glover E, Pennock D M, Flake G and Nielsen F, Persistence of web references in scientific research, *IEEE Computer*, 34 (2) (2001) 26-31.

3 Rumsey M, Runaway train: Problems of permanence, accessibility, and stability in the use of web sources in law review citations, *Law Library Journal*, 94 (2) (2002) 27-39.

4 Veena B and Sampath Kumar B T, Web citation behaviour in scholarly electronic journals in the field of library and information science, *Webology,* 5 (2) 2008. Available at: www.webology.ir/2008/v5n2/a57.html (Accessed 13 May, 2013).

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

6 Bar-Ilan J and Peritz B C, The lifespan of Informetrics on the web: An eight year study (1998-2006), *Scientometrics,* 79 (1) (2009) 7-25.

7 Sampath Kumar B T and Prithviraj K R, Availability and persistence of web citations in Indian LIS literature, *The Electronic Library*, 30 (1) (2012) 19–32.

8 Lawrence S, Online or invisible, *Nature,* 411 (6837) (2001) 521.

9 Casserly M F and Bird J E, Web citation availability: analysis and implications for scholarship, *College and Research Libraries*, 64 (7) (2003) 300-317.

10 Hester E J, Heilig L F, Drake A L, Johnson K R, Vu C T, Schilling L M and Dellavalle R P, Internet citations in oncology journals: a vanishing resource, *Journal of the National Cancer Institute*, 96 (12) (2004) 969-971.

11 Sellitto C, A study of missing web-cites in scholarly articles: towards an evaluation framework, *Journal of Information Science*, 30 (6) (2004) 484-495.

12 Patil S K, Chikate R V, Wadekar P P and Joshi S, Measuring Impact of Web Sources in Conference Proceedings: A Citation Analysis, *In proceedings of Naclin.* (2007) 142-153.

13 Janakiramaiah M and Doraswamy M, Measuring Impact of Web Resources in Conference Proceedings: A Citation Analysis, *International Caliber,*2011. 541-549.

14 Riahinia N, Zandian F and Azimi A, Web citation persistence over time: a retrospective study, *The Electronic Library,* 29 (5) (2011) 609-620.

15 Nagaraja A, Joseph S A, Polen H H and Clauson K A, Disappearing act: persistence and attrition of uniform resource locators (URLs) in an open access medical journal, *Program: electronic library and information systems*, 45 (1) (2011) 98-106.

16 Saberi M K and Abedi H, Accessibility and decay of web citations in five open access ISI journals. *Internet Research*, 22 (2) (2012) 234–247.

17 Sampath Kumar B T and Manoj Kumar K S, Decay and half-life period of online citations cited in open access journals, *The International Information & Library Review,* 44 (4) (2012a) 202–211.

18 Sampath Kumar B T and Vinay Kumar D, HTTP 404-page (not) found: Recovery of decayed URL citations, *Journal of Informetrics,* 7 (2013) 145–157.

19 Harter S P and Kim J H, Electronic journals and scholarly communication: A citation and reference study, *Information Research,* 2 (1996).

20 Davis P M and Cohen S A, The effect of the web on undergraduate citation behaviour 1996-1999, *Journal of the American Society for Information Science and Technology*, 52 (4) (2001) 309-314.

21 Herring D H, Use of Electronic resources in scholarly electronic journals: A citation analysis, *College & Research Libraries,* 63 (4) (2002) 334-340.

22 Kushkowski J D, Web citation by graduate students: a comparison of print and electronic theses, *Portal: Libraries and the Academy,* 5 (2) (2005) 259-276.

23 Olfson E and Lawrence J, Accessibility and longevity of internet citations in a clinical AIDS journal, *AIDS Patient Care STDS,* 19 (1) (2005) 5-8.

24 Megnigbeto E, Internet-based resources citation by undergraduate students: A case study of library and information science students in Benin, *The International Information & Library Review*, 38 (2) (2006) 49–55.

25 Ducut, E, Liu F and Fontelo P, An update on Uniform Resource Locator (URL) decay in MEDLINE abstracts and measures for its mitigation, *BMC Medical Informatics and Decision Making*, (2008) doi:10.1186/1472-6947-8-23.

26 Chuanfu Chen, Kai Sun, Gang Wu, Qiong Tang, Jian Qin, Kuei Chiu, Yushuang Fu, Xiaofang Wang, and Jing Liu, The impact of internet resources on scholarly communication: A citation analysis, *Scientometrics,* 81 (2) (2009) 459–474.

27 Yang S, Qiu J and Xiong Z, An empirical study on the utilization of web academic resources in humanities and social sciences based on web citations, *Scientometrics*, 84 (1) (2010).

28 Saberi M K, Isfandyari-Moghaddam A and Mohammad Esmaeil S, Web Citations Analysis of the JASSS: the First Ten Years, *Journal of Artificial Societies and Social Simulation,* 14 (4) (2011) 22.

29 Mardani A, An investigation of the web citations in Iran's chemistry articles in SCI, *Library Review*, 61 (1) (2012) 18–29.

30 Sadat-Moosavi A, Isfandyari-Moghaddam A and Tajeddini O, Accessibility of online resources cited in scholarly LIS journals: A study of Emerald ISI-ranked journals, *Aslib Proceedings*, 64 (2) (2012) 178-192.

31 Sampath Kumar B T and Manoj Kumar K S, Persistence and Half-life of URL citations cited in LIS Open Access Journals, *Aslib Proceedings*, 64 (4) (2012b) 405 – 422.

32 Jalalifard M, Norouzi Y and Isfandyari-Moghaddam A, Analyzing web citations availability and half-life in medical journals: A case study in an Iranian university, *Aslib Proceedings*, 65 (3) (2013) 242 – 261.