DESIDOC Journal of Library & Information Technology, Vol. 32, No. 4, July 2012, pp. 347-352

Web Impact Factor of Select National Libraries' Websites

Paramjeet K. Walia and Monika Gupta

Department of Library and Information Science, University of Delhi, Delhi-110 007
E-mail: pkwalia2002@gmail.com; monikagupta38@gmail.com

ABSTRACT

The objective of this study is to examine the linking on websites of national libraries and to find out their web impact factor and amount of information present on these websites in the form of rich files. The methodology which is followed in this study is basically investigative in nature. For the purpose of analysing the link structure, data is collected in the month of December 2009, January 2010, April 2010, and May 2010. The research revealed that among the selected national libraries, websites of national libraries' of America, Australia and Britain were more visible and hosted the more content compare to the websites of India, Namibia, and South Africa. A short survey is conducted to find the number of functional national libraries websites in the world. Among the 163 countries (which have national libraries) 106 countries have the websites. In this study two national libraries are selected from each continents of the world except American continent.

Keywords: National library, website, webometrics, web impact factor

1. INTRODUCTION

Libraries form a vital part of the world's system of education and information storage and retrieval. National libraries advance the achievement of national goals, economic development and adequate resources, through the provision of relevant and timely information. They promote and develop the nation's cultural heritage, promote literacy, and achieve national goals in the field of library and information services. Nowadays, via ICT-assisted dynamic environment, national libraries have very suitable facilities for playing their role. Undoubtedly, one powerful website is adequate for meeting users' needs of national libraries in anywhere. As the world wide web becomes a global tool of information, national libraries have a great chance of contributing much valuable information through their websites. With libraries shifting their role from being custodians of collection-based traditional information resources to being providers of access-based digital information resources, the library websites have assumed importance.

Libraries have to disseminate and facilitate access to variety of information to users through their websites. To succeed, however, the sites must be well-organised, informative, easy to navigate, and accessible in both native and foreign languages. National libraries need to make necessary investments for introducing and making accessible their website content. For this, the national libraries should invest on designing and managing of websites.

2. OBJECTIVES OF STUDY

The national library has a unique role to play in meeting the information needs of all citizens of a country. It is a custodian and provider of the nation's key knowledge resources. Many national libraries have their own websites to provide a view of their collection, services and activities but, despite that, there has been relatively little analysis on the visibility on the web and linking with other webpages. This brings us to an unavoidable stand wherein an analysis of the linking and visibility of these websites has become extremely vital. The present study was carried out to examine the linking on websites of national libraries and to find out their web impact factor (WIF) and amount of information present on these websites in the form of rich files.

3. SCOPE AND LIMITATION

According to IFLA[1] (International Federation of Library Associations and Institutions), there are 163 national libraries including those claiming function but not having the actual title. In this study two national libraries are selected from each continent of the world except

Received 14 October 2011, accepted 30 January 2012, online published 18 July 2012

American continent because most of the websites of national libraries are available in their national language only. For this study nine websites of national libraries are selected. The libraries covers in this study are as shown in Table 1.

4. METHODOLOGY

The methodology followed is basically investigative in nature. The methodologies followed at the different stages of the research to draw the conclusions are review of literature, online searching, and calculation of WIF. For the purpose of analysing the link structure, data is collected in the month of December 2009, January 2010, April 2010 and May 2010. After collecting the data of two months, a gap is given for two month to observe the variation and after that again data is collected for another two months. The purpose of this is to find out trend of growth of webpages and in-links.

5. WEBOMETRICS

The term webometrics was first coined by Almind & Ingwersen[1]. The science of webometrics tries to measure the world wide web to get knowledge about the number and types of hyperlinks, structure of the world wide web and usage patterns. According to Bjorneborn and Ingwersen the definition of webometrics is the study of the quantitative aspects of the construction and use of information resources, structures and technologies on the web drawing on bibliometric and informetric approaches'[2]. A second definition of webometrics has also been introduced by Thelwall[3] as 'the study of web-based content with primarily quantitative methods for social science research goals using techniques that are not specific to one field of study'[3]. Similar scientific fields are: Bibliometrics, informetrics, scientometrics, and web mining.

5.1 Web Impact Factor (WIF)

The idea of measuring WIF as one of the quantitative indicators (or the average link frequencies) was developed in 1998 by Peter Ingwersen. The WIF is a 'snapshot' of a search engine database at a specific time. The WIF provides quantitative tools for ranking, evaluating, categorising, and comparing websites, top-level domains and sub-domains. Compared with the content of a journal paper, the content of a web resource lacks peer review and thus lacks quality control. The WIF is therefore not exactly the equivalent of the journal impact factor (JIF). However, the WIF was inspired by the JIF. There are three types of links:

(a) Outlinks: Outgoing links from webpages

(b) Inlinks/backlinks: Links coming into a site from other sites

(c) Self-links: Links within the same site (from one page to another page)

There are three types of WIF:

(a) Overall WIF

(b) Inlink (revised) WIF

(c) Self-link WIF

For calculation of the overall WIF of a website, the numerator is the number of inlink page from outside the site and self-link pages within the site. For the inlink (or revised) WIF, the numerator is the number of inlink pages counted from outside the site. And for self- link WIF, numerator is number of self-link pages counted from within site. In all these cases, the denominator remains the same, i.e., the number of webpages within the website. The calculation for WIF is:

WIF= A/D

A= Total link pages (all inlink and self-link pages);

D= Number of webpages published in the website which are indexed by the search engine, not all webpages available in the website:

The calculation for revised WIF to exclude self-links is:

Revised WIF= (B/D)

Table 1. National libraries covered

Name of National Library	Abbreviation Used	Website Address
The British Library	BL	http://portico.bl.uk/
German National Library	GNL	http://www.d-nb.de/eng/index.htm
Library of Congress	LC	http://www.loc.gov/index.html
National Library of Australia	NLA	http://www.nla.gov.au/
National Library of China	NLC	http://www.nlc.gov.cn
National Library, India	NLI	http://www.nationallibrary.gov.in
National Library of Namibia	NLN	http://www.nln.gov.na
National Library of New Zealand	NLNZ	http://www.natlib.govt.nz/
National Library of South Africa	NLSA	http://www.nlsa.ac.za/NLSA/

A= Total links to a website (all inliks and self-link pages);

B= Inlinks to the website (this is a subset of A);

C=Self-link and navigational links within the same website;

D= number of webpages published in the website which are indexed by the search engine, not all webpages available in the website[4].

WIF is a form of measurement used to determine the relative standing of websites in particular fields, or a country; for instance, academic websites in a country. The higher the impact factor, higher is the perceived reputation of the website. The WIF answers the question 'what impact has this website had?' A WIF is a measure of the frequency with which the average webpage in a website has been linked at a given point of time. In general, a website with a higher impact factor may be considered to be more prestigious or of a higher quality than those websites with a lower impact factor.

6. DATA ANALYSIS

For the purpose of analysing the link structure, the data is collected in the month of December 2009, January 2010, April 2010 and May 2010. After collecting the data of two months, a gap is given for two month to observe the variation and after that again data is colleted for another two months. The purpose of this is to find out trend of growth of webpages and inlinks.

6.1 Analyses of Web Links

These analyses require two types of data, i.e., number of webpages on each website and number of inlinks on a website. In the given table, the data is presented with respect to the month in which it is collected (Table 2).

The LC has the highest number of webpages during the month of December and January but in the next two months, i.e., April and May, the NLA has the maximum number of webpages. The NLI has the lowest number of webpages in all the four months during which the data is collected. Number of webpages increased on the website of BL in all the four months. On the other hand webpages are increased in the first two months on the websites of GNL, LC, and NLSA and after two months gap when again data is collected the webpages are decreased in April and in May again it increased. The NLN website has minimum number of change in number of webpages. Number of webpages increased and decreased over a time shows, that websites are regularly maintained and updated. Inlinks data reveals that among all the libraries, LC has the highest number of inlinks during all the months in which data is collected (Table 3). NLN has the lowest number of inlinks in all the four months during which the data is collected. Number of inlinks are changing on all the websites, they are not static, which shows that the websites are regularly loaded by current information and outdated information is removed after a time period.

6.2 Calculation of Web Impact Factor

The calculation of the WIF requires two types of data, namely, number of webpages and number of inlinks of a website. Both these data have already been given in the former tables. Thus, calculation of WIF is done, for which, the calculated values are displayed in the Table 4. This data is also presented in a month-wise manner, i.e., for each month a separate WIF value was calculated. The last column in Table 4 gives the average of the four months for easy ranking.

Table 4 reveals that the LC has the highest WIF in all the four months. It leads with an average WIF of 18.609 followed by GNL with 14.618, then NLNZ with 14.48, then BL with 5.268, then NLC with 4.984, then NLI with 4.669, and then NLA with 3.929, NLN with 2.156 and the last rank is received by NLSA with 0.533.

6.3 Information about Rich Files

The rich files constitute of files like portable document format (.pdf), Microsoft Word (.doc) and Microsoft Power

Table 2. Number of webpages

Name of library	Webpages				Average
	Dec-09	Jan-10	Apr-10	May-10	
BL	139302	153601	201518	201690	174027.75
GNL	2107	2124	2058	2121	2102.5
LC	271499	272870	268912	476769	322512.5
NLA	143579	152930	1150041	1123265	392453.75
NLC	31192	32592	28807	26380	29742.75
NLI	51	52	58	38	49.75
NLN	86	86	85	85	66.25
NLNZ	10286	9897	7742	8632	9139.25
NLSA	2646	2728	2088	2688	2537.5

Table 3. Number of inlinks

Name of library	Inlinks					
	Dec-09	Jan-10	Apr-10	May-10	Average	Points
BL	1121433	865197	893753	596361	869186	3
GNL	23432	22943	36681	39728	30696	1
LC	4567409	4386143	7465337	6570502	5747347.75	4
NLA	901099	1141485	1265859	988199	871160.5	3
NLC	135684	132082	139321	176757	145961	2
NLI	214	220	274	210	229.5	1
NLN	169	182	190	196	184.25	1
NLNZ	117466	123571	140748	136734	129629.75	2
NLSA	1506	1623	1160	1120	1352.25	1

Table 4. Calculation of web impact factor (WIF)

Name of library	Dec-09 (*InL/WP*)	Jan-10 (*InL/WP*)	Apr-10 (*InL/WP*)	May-10 (*InL/WP*)	Average
BL	08.05	05.632	04.435	02.956	05.268
GNL	11.121	10.801	17.823	18.730	14.618
LC	16.822	16.074	27.761	13.781	18.609
NLA	06.275	07.464	01.100	00.879	03.929
NLC	04.349	04.052	04.836	06.700	04.984
NLI	04.196	04.23	04.724	05.526	04.669
NLN	01.965	02.119	02.235	02.305	02.156
NLNZ	11.419	12.485	18.179	15.840	14.480
NLSA	0.569	00.594	00.555	00.416	00.533

Point (.ppt). These files represent the amount of published information on a website. The complete, file type-wise data regarding number of rich files of national libraries has been displayed in the next table in a month-wise manner (Table 5).

Month wise, NLA is at rank number one in all the four months of data collection. LC takes the second rank in all the four months, followed by British Library taken third rank. National Library of Namibia which comes at the last rank. The analyses of data reveal that the number of .pdf files exceeds the number of .doc as well as .ppt files. Also the numbers of .ppt files are not sufficient on almost all the websites of national libraries covered under the study.

7. FINDINGS

The major findings of the study are:

(i) Webpages and Inlinks

The rank list of websites of national libraries on the basis of maximum number of webpages has been shown in Fig. 1.

Figure 1 shows that on average, NLA hosts the highest number of webpages on the web followed by LC and BL. It is NLI India that hosts the lowest number of webpages.

Figure 2 gives number of inlinks of websites of national libraries. It shows that LC has the maximum number of inlinks, i.e., 5747347.75 followed by NLA, whereas NLN takes the last position with an average number of 184.25 inlinks.

(ii) Web Impact Factor

According to the data provided detailed calculation of WIF is done and values are displayed in Fig. 3.

Figure 3 shows that LC rank number one in its impact on the web with far more the number of inlinks as compare to the number of webpages on its website. It leads with a WIF of 18.609 followed by the GNL at the second position with a WIF of 14.618 while National Library of Australia (NLA) with a WIF of 3.929, comes at number seven, though it stands number one in the rank list of webpages and number two in the rank list of inlinks.

(iii) Rich Files

The rich text files constitute of files like .pdf, .doc, .ppt. The rank list on the basis of rich files is shown in Fig. 4 that shows that NLA hosts the highest number of rich files on its website, i.e., 27417.5 on an average followed by LC with 12136.5 on an average. It is the website of NLN which comes at the last in rank list with an average 4.75 files.

Table 5. Number of rich files of National Library's websites

Name of library	Dec-09				Jan-09			
	.pdf	.doc	.ppt	Total	.pdf	.doc	.ppt	Total
BL	2180	25	14	2219	2210	22	13	2245
GNL	675	0	0	675	756	0	0	756
LC	14000	377	466	14843	13000	393	487	13880
NLA	24600	1170	427	26197	22400	1170	438	24008
NLC	1070	97	5	1172	1040	103	4	1147
NLI	23	0	0	23	24	0	0	24
NLN	4	1	0	5	4	1	0	5
NLNZ	252	409	35	696	262	431	37	730
NLSA	117	1	0	118	120	1	0	121

Apr-10				May-10				Grand Total	Average	Points
.pdf	.doc	.ppt	Total	.pdf	.doc	.ppt	Total			
1610	18	24	1652	1985	24	31	2040	8156	2039	2
699	0	0	699	725	0	0	725	2855	713.75	1
8810	379	416	9605	9410	368	440	10218	48546	12136.5	4
28300	895	346	29541	28546	900	478	29924	109670	27417.5	4
906	96	4	1006	960	95	4	1059	4384	1096	1
28	0	0	28	28	0	0	28	103	25.75	1
4	1	0	5	4	0	0	4	19	4.75	1
267	432	62	761	277	432	60	769	2956	739	1
120	0	1	121	120	0	1	121	481	120.25	1

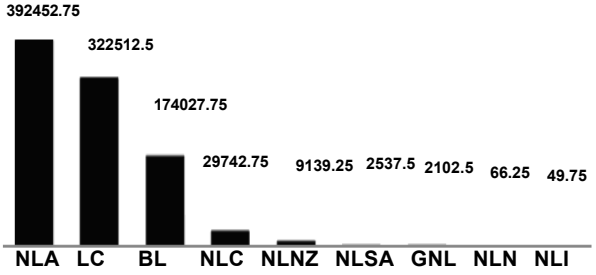

Figure 1. Rank list on the basis of number of webpages.

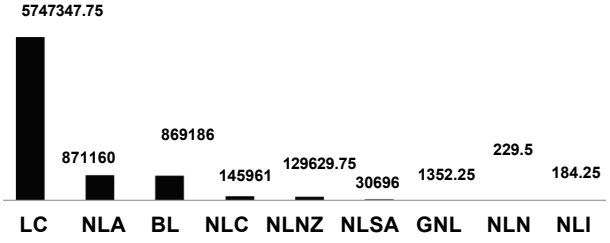

Figure 2. Rank list on the basis of number of inlinks.

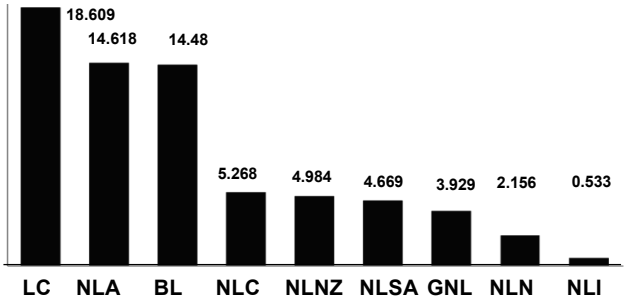

Figure 3. Rank list on the basis of web impact factor.

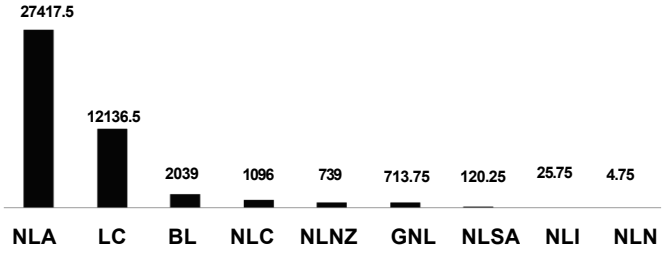

Figure 4. Rank list on the basis of rich files.

8. SUGGESTIONS

(a) The websites of some libraries like NLI and NLN having minimum number of webpages, should go to a deeper extent to make maximum number of its webpages indexed by the search engines, so that it becomes easier for the surfer to get access to the information available on the website.

(b) In addition to providing maximum number of documents, the libraries should also try to provide information in different formats for the ease of the users. None of website under study hosts an equal number of types of files covered here.

(c) For receiving maximum number of hits from users national libraries should increase the information on their websites.

9. CONCLUSIONS

Website makes strong impact on the image of library, information centre, and documentation centre. It is very important when information providers are more concerned about exploring its resources world wide like national libraries. From the study, it can be concluded that LC is more visible websites among the other websites selected. The LC has the maximum number of WIF. The NLA ranked in the first position with respect to number of webpages among the various selected national libraries. The LC leads all the libraries with maximum webpages in the first two month than NLA hold the top position in the later two months in the number of webpages. The LC had the maximum number of inlinks, i.e., 5747347.75 followed by NLA. The website of NLA holds the maximum number of rich content files on its website.

REFERENCE

1. IFLA. National libraries of the world: Address list. http://ifla.queenslibrary.org/VI/2/p2/national-libraries.htm (accessed on 11 July 2009).

2. Bjorneborn, Lennart & Ingwersen, Peter. Toward a basic framework for webometrics. *JASIST*, 2004, **55**(14), 1216-227.

3. Mike, Thelwall. Introduction to webometrics: Quantitative web research for the social sciences. Morgan & Claypool, San Rafael, 2009.

4. Li, Xuemei. A review of the development and application of the web impact factor. *Online Inf. Rev.,* 2003, **27**(6), 407-17.

About the Authors

Dr Paramjeet K. Walia is Head and Associate Professor in the Department of Library and Information Science, University of Delhi. She did her MLISc, MPhil and PhD on Government Documents from DU. She also taught for 14 years at Panjab University, Chandigarh. Her areas of interest are: Public library system, education for LIS, government documents, information systems and programmes and library management. She has contributed several papers in Journals and conference/ seminars of national and international repute. She is Editor of *Journal of Library and Information Science*. She had also served in the National Medical Library and Delhi Stock Exchange Library in India.

Ms Monika Gupta is Research Scholar in the Department of Library and Information Science, University of Delhi. She did her BLISc, MLISc and MPhil from University of Delhi. She has also worked as Assistant Librarian in G.B. Pant Engineering College, Government of NCT of Delhi. She has contributed several papers in the journals and conference of national and international repute.