CITING PATTERNS IN OPEN ACCESS JOURNALS: A STUDY OF D-LIB MAGAZINE

Moin Ahmad, *American University of Ras A L Khaimah, Ras A L Khaimah, UAE*
Mohammad Nazim, *Banaras Hindu University, Varanasi, UP*

Abstract

The purpose of this study was to investigate the pattern of citing references of research articles published in D-Lib Magazine during 2002 to 2008. A total of 4775 citations were collected from 295 articles published during 2002 to 2008. Articles classified as editorial materials, power point slides, book reviews, columns, reports and news items were not considered for the analysis. References of each article were collected and Microsoft Office Excel 2007 was used for analyses. The various analyses focus on year-wise distribution of articles and cited references, types of documents cited, country and language of cited documents, file format and domain of cited references, etc. The study shows the changing trends of research in the field of library & information science in the field of digital libraries particularly with the introduction of Internet and World Wide Web. This change can be seen in digital library research as researchers have been used digital and web resources to conduct their research.

Keywords: *Citation Analysis, Open Access Journals, D-Lib Magazine, Citation Pattern.*

Introduction

Citation analysis is a sub-area of bibliometrics. Bibliometrics is the application of quantitative methods to the study of communication media such as books and published articles (Beker, 1990). Citation analysis is the examination of the frequency, patterns and graphs of citations in scholarly communication. It uses citations in scholarly works to establish links to other works or other researchers. With the increasing involvement of technology in the scholarly communication process, the style of traditional citing reference has changed (Mukherjee, 2009). In scholarly texts, authors are now using Internet based electronic resources and include them in reference list to cite sources they have consulted. Citation pattern and journal ranking have been examined for a long time. Citation analysis has been used extensive to determine and define core journal collection in academic libraries. It also provides acquaintance about what part of literature is cited most, how long the literature remains useful to readers and languages of most cited publications and useful for collection development policies, individual item selection and retention and binding decisions. Therefore, there is a need to understand the pattern of citations to OA journals' articles.

Digital library research is an emerging area in the field of library and information Science. D-Lib Magazine is an important open access journal that reports the results of research in the area of Digital Library.

Objectives of the Study

The present study is an effort to investigate the citing patterns of the articles published in D-Lib Magazine. The magazine is currently published six times a year. The full contents of the magazine, including all back issues, are available free of charge at the D-Lib web site (http://www.dlib.org). The magazine is also accessible through the Directory of Open Access Journals (DOAJ, http://www.doaj.org/) and Open J-gate (http://www.openj-gate.com).

The motivation for writing this paper is to answer the following research questions:
- What is the pattern of citing references of open access journals?
- Do authors refer more to electronic information sources than to print information sources in research papers for open access journals?
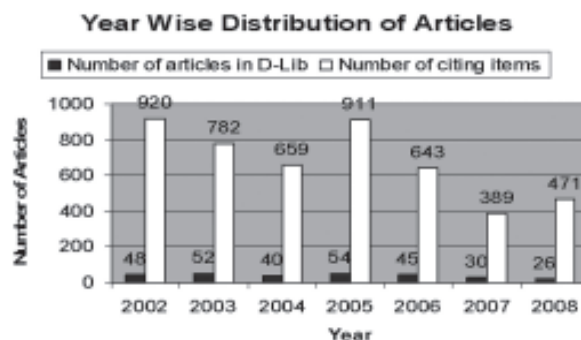
Methods

The purpose of this study was to investigate the pattern of citing references of research articles published in open access journals and therefore, D-Lib Magazine was selected as source for collecting references of the articles for citation analysis. All issues of D-Lib Magazine from 2002 to 2008 were consulted. Articles classified as editorial materials, power point slides, book reviews, columns, reports and news items were not considered for the analysis. References of each article were collected and Microsoft Office Excel 2007 was used for analyses. A total of 4775 citations were collected from 295 articles published during the sample period with an average of 16.18 citations per articles.

Analysis

Year-Wise Distribution of Articles and Cited References

A total of 4775 citations were collected from 295 articles published during 2002-2008, thus an average of 16.18 citations per article. Figure 1 shows the year-wise distribution of number of articles and citations. The highest number of citations was found in the year 2002 with an average of 19.16 citations each article and nearly equal number of citations in 2005 with an average of 16.87 citations each article. The lowest number of citations was counted in the year 2007 with the average of 12.96 citations per article.

Figure-1

Types of Documents

The study regarding form-wise distribution of citations have been done to identify the most dominant form of documents cited in the D-Lib Magazine. Table 1 shows that the types of documents cited in the articles of D-Lib Magazine. It is to be observed from the table that although various categories of documents are cited by the authors, the main forms of the documents cited in the articles of D-Lib Magazine are web resources (56.84%) followed by journal articles (24.40%), books (8.06%), conference proceedings (6.66%), workshop reports (1.76%), research reports (1.57%), Symposium (0.54%) and thesis (0.17%).

It is important to note that with the increasing involvement of technology in the scholarly communication process, the style of traditional cited reference has changed. In scholarly communication, authors are now using web resources/hyperlinks either inside the text or in reference list to site the sources they have consulted. Present study shows this pattern clearly as more than 50% citations are web resources.
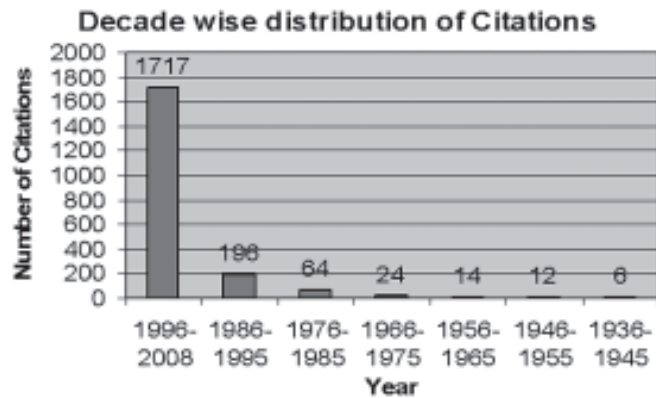
Table 1 Types of Documents

S. No.	Form of Document	Rank	2002	2003	2004	2005	2006	2007	2008	Total	%
1	Web resources	1	499	396	415	607	346	191	260	**2714**	56.84
2	Journal Articles	2	210	220	124	177	173	137	124	**1165**	24.40
3	Books	3	98	105	55	39	36	23	29	**385**	8.06
4	Conference Proceedings	4	76	37	45	50	61	22	27	**318**	6.66
5	Workshop Repots	5	15	10	13	20	13	5	8	**84**	1.76
6	Research reports	6	15	6	4	9	13	8	20	**75**	1.57
7	Symposium	7	6	6	3	4	1	3	3	**26**	0.54
8	Thesis	8	1	2		5		0		**8**	0.17
	Total		920	782	659	911	643	389	471	4775	

Year-Wise Distribution of Cited Documents

In order to identify the year-wise distribution of cited documents, the year of publication of each cited documents was recorded for the analysis. Out of total 4775 citations, the year of 2442 cited documents could not be traced out. Thus, 2033 citations which mentioned the year of publication were analyzed. It observed from figure 2 that 84.45% cited documents were published during 1996-2006 and remaining 15.55% were published during 1936-1995. It observed that most of the authors used latest sources of information in scholarly communication.

Figure-2

Decade wise distribution of Citations

Distribution of Citations by Country

For the analyses of geographical scattering of cited references, country in which each cited documents published was recorded. Only journal citations were included in the study because country of other types of cited documents could not be traced. The analysis reveals that the cited documents used in the research articles of D-Lib Magazine emanated from different countries. As mentioned in table 2 highest number of citations was from United States (867) followed by United Kingdom (196) and Germany (42).

Table 2 Distribution of Citations by Country

S. No.	Country	Rank	Frequency	%
1	US	1	867	74.42
2	UK	2	196	16.82
3	Germany	3	42	3.61
4	Canada	4	14	1.20
5	Netherland	4	14	1.20
6	china	5	5	0.43
7	Australia	6	4	0.34
8	Brazil	7	3	0.26
9	Switzerland	7	3	0.26
10	India	8	2	0.17
11	Austria	8	2	0.17
12	France	8	2	0.17
13	Israel	8	2	0.17
14	Italy	8	2	0.17

Cont...

15	Singapore	8	2	0.17
16	Hungry	9	1	0.09
17	Japan	9	1	0.09
18	Newziland	9	1	0.09
19	Sweden	9	1	0.09
20	Virginia	9	1	0.09
	Total		***1165**	

* Only journal articles were included for this analysis

Language-Wise Distribution of Cited References

The study also identified the language of the documents cited in D-Lib Magazine. It may be observed from table 3 that documents cited in D-Lib Magazine were published in various languages. However, English becomes the most prominent language of the cited documents. Out of total 1165 documents 1054 documents was in English language. English is followed by German, Chinese and French with 38, 15 and 7 documents respectively.

Table 3 Language-Wise Distribution of Citations

S. No.	Language	No. of articles
1	English	1054
2	German	38
3	Chinese	15
4	French	7
5	Italian	5
6	Japanies	3
7	Spanis	3
8	Portuguese	2
9	Swidish	2
10	Hungrian	2
11	Other Languages	34
	Total	**1165**

66. Only journal articles were included for this analysis

Composition of Web Citations by Types of File Format

Web citations were found in various types of file format, such as audio, video, image and presentations, etc. In order to analyze the different file formats of the cited web references in the present study, the web citations had been categorized into seven categories:

- HTML (Hyper Text Markup Language): web documents created in HTML scripting language;
- PDF (Portable Document Format): the file format for documents created using Adobe Acrobat;
- PPT (Power Point Presentations);
- DOC (Documents created using MS-Word);
- RTF (Rich Text Formats): a text file format that includes formatting features, such as bold, italic, and underline text;
- HTM;
- Others (Unable to identify the file format).

As table 4 reveals that the greatest numbers of cited web references are in HTML files (533) followed by PDF files (339) HTM files (216) and DOC files (23) The file format of 1594 web citations could not be traced and therefore, grouped into the other category.

Table 4 Composition of Web Citations by Type of File Formats

S. No.	Types of the files	Frequency	%
1	HTML	533	19.64
2	PDF	339	12.49
6	HTM	216	7.96
4	DOC	23	0.85
3	PPT	6	0.22
5	RTF	3	0.11
7	Other	1594	58.73
	Total		**2714**

Composition of URLS of the Cited References by Type of Domain

The URL is an address of a digital document on the web. A URL essentially has four parts: Protocol, domain, directory and file. A domain name is a way to identify and locate computers connected to the Internet. Two organizations cannot have the same domain name. A domain name always contains two or more components separated by periods, which are called "dots." Some examples of domain names are: yahoo.com, oclc.org, tcs.co.in and nasa.gov. A domain name can often tell the user if it is a government site, an academic site or a commercial site. Some common top-level domain name endings are:

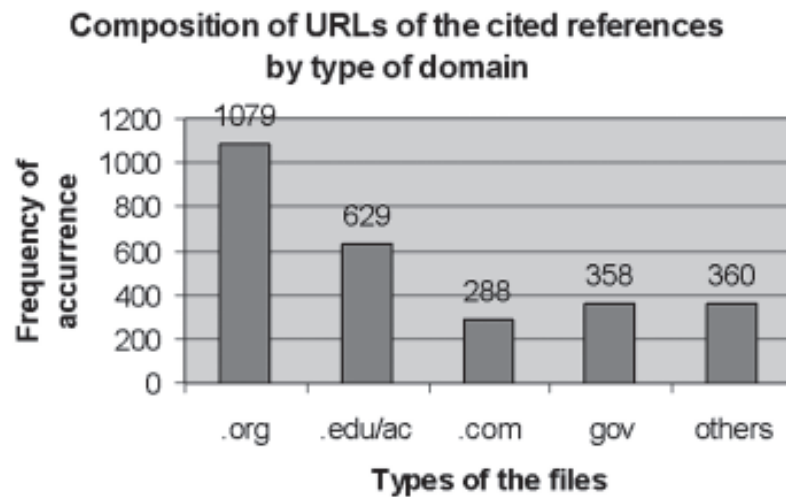- .com or .co=a commercial organization;
- .edu or .ac=an educational organization;
- .gov=an official government site;
- .org=mostly non-profit organizations;
- .net=traditionally it was for network organizations, but now can be used by anyone.

Country codes: A Country code is a two-letter abbreviation for a particular country e.g. "in" for India, "uk" for United Kingdom and "us" for United State, etc.

Here we listed only five types of domains and remaining domains grouped into others category.

They are .org, .edu/.ac, .com and .gov while those domains not falling into any of these categories fall into the "other" category.

Figure 3 show the domains of URLs of the cited documents. The highest number of URLs of cited documents come under .org (1079) and .edu/ac(629) domains. This indicates that most of the web citations in the present study are available through website of various non-profit organizations and educational institutions and they are more scholarly in nature and the likely sustainability of such resources is better than that of .com sites.

Figure-3

Composition of URLs of the cited references by type of domain

Ranking of Cited URLs

The Internet is one of the largest and most complete tools for information exchange ever made available to the global population. Therefore, the present study is to judge the variable and productive web sites cited in the D-Lib Magazine.

The 37 most cited URLs were listed ranked from greater to lesser number of citations. It may be observed from table 5 that the www.loc.gov is the most citied URL with 4.61% citations; followed by www.w3.org, www.openarchives.org, www.dublincore.org, www.jisc.ac.uk, and www.oclc.org with 2.58%, 2.32%, 2.06%, 2.03% and 1.99% citations respectively.

Table 5 Rank list of URLs

S. No	URL	Rank	Frequency	%
1	www.loc.gov.	1	125	4.61
2	www.w3.org	2	70	2.58
3	www.openarchives.org	3	63	2.32
4	www.dublincore.org	4	56	2.06
5	www.jisc.ac.uk	5	55	2.03
6	www.oclc.org	6	54	1.99
7	www.dlese.org	7	34	1.25
8	www.arl.org	7	34	1.25
9	www.ifla.org	8	29	1.07
10	www.niso.org	9	26	0.96
11	www.doi.org	10	21	0.77
12	www.nla.gov.au	11	20	0.74
13	www.apsr.edu.au	11	20	0.74
14	www.rlg.org	12	19	0.70
15	http.serc.carleton.edu	13	17	0.63
16	www.cni.org.	14	15	0.55
17	www.surf.nl	15	14	0.52
18	www.digitalpreservation.gov	16	13	0.48
19	www.diglib.org	16	13	0.48
20	www.steve.museum	16	13	0.48
21	http.arxiv.org	16	13	0.48
22	http.en.wikipedia.org	17	12	0.44
23	http.hdl.handle.net	17	12	0.44
24	www.language-archives.org	18	11	0.41
25	www.leeds.ac.uk	18	11	0.41
26	www.rdn.ac.uk	18	11	0.41
27	www.ukoln.ac.uk	18	11	0.41
28	http.news.bbc.co.uk	19	10	0.37
29	www.cnri.reston.va.us	19	10	0.37
30	www.ecs.soton.ac.uk	19	10	0.37
31	www.archive.org	19	10	0.37
32	www.clir.org	19	10	0.37
33	http.alpha.fdu.edu	20	9	0.33
34	http.crs.dlese.org	20	9	0.33
35	www.alexandria.ucsb.edu	20	9	0.33
36	www.openurl.info	20	9	0.33
37	http.purl.org	21	8	0.29
38	www.archimuse.com	21	8	0.29
39	www.ariadne.ac.uk	21	8	0.29
40	www.library.gatech.edu	21	8	0.29
41	URL with 1-7 Citations	1804	66.47	
Total		**2714**		

Authorship Pattern

Authorship pattern of the cited journal articles was analyzed to know the trends of research. Only journal articles were included for this analysis.

The number of authors contributing to each cited articles range from one to eighteen. However, 47.81% cited articles contributed by single author while 27.73% articles contributed by two authors. Articles with more than three authors account for 24.46% of total number of cited journal articles which clearly indicates that collaborative efforts are common in Digital library research. Contribution to each cited article by number of authors is shown in detail in table 6.

Table 6 Authorship Pattern

Sl. No.	Year	Single Author	Two Author	Three Author	Four Author	Five Author	More Than Five Author	Total
1	2002	113	54	23	8	5	7	**210**
2	2003	110	59	36	6	3	4	**218**
3	2004	65	34	13	7	2	4	**125**
4	2005	89	42	22	15	6	4	**178**
5	2006	65	41	22	17	10	18	**173**
6	2007	63	42	10	15	2	5	**137**
7	2008	52	51	11	4	1	5	**124**
Total	**557**	**323**	**137**	**72**	**29**	**47**	**1165**	
%	47.81	27.73	24.46	8.58	2.49	4.34		

*Only cited journal articles were considered for this analysis

Ranking of Cited Journals

Journals are very useful for researchers for the scientific communication but their increasing cost puts the librarian to study the quality, usefulness and suitability to a particular group of users. Therefore, the present study is to judge the variable and fruitful journals, cited by D-Lib Magazine.

1165 citations analyzed in the present study were published in 297 journals. Table 7 listed 49 most cited journals. It may be observed that D-Lib Magazine is the most cited journal as it was cited 341 times in D-Lib Magazine. Journal of the American Society of Information Science, Communications of the ACM, Ariadne, and Library Journal occupies the 2nd , 3rd , 4th and 5th place respectively in the rank list of most cited journals.

Author Productivity Table-7: Ranking Of Journals

S. No.	Name of Periodical	Rank	Country	Frequency	%
1	D-Lib Magazine	1	U S	341	29.27
2	Journal of the American Society for Information Science and Technology	2	U S	50	4.29
3	Communications of the ACM	3	U S	37	3.18
4	Ariadne	4	Germany	25	2.15
5	Library Trends	5	U S	21	1.80
6	RLG DigiNews	6	U S	20	1.72
7	International Journal on Digital Libraries	7	U S	18	1.55
8	Library Hi Tech	8	GB	17	1.46
9	Library Journal	9	U S	16	1.37
10	Information Processing & Management	9	GB	16	1.37
11	College & Research Libraries	10	U S	14	1.20
12	Journal of Information Science	11	GB	13	1.12
13	Serials Review	12	GB	12	1.03
14	Learned Publishing	12	GB	12	1.03
15	Practical Assessment Research and Evaluation	13	U S	11	0.94
16	First Monday	13	U S	11	0.94
17	Journal of Digital Information	14	GB	9	0.77
18	Serials	15	GB	8	0.69
19	Computers in Libraries	15	U S	8	0.69
20	Portal: Libraries and the Academy	16	U S	7	0.60
21	Journal of Documentation	16	GB	7	0.60
22	Journal of Academic Librarianship	16	GB	7	0.60
23	Information Technology & Libraries	16	U S	7	0.60
24	Reference Services Review	17	GB	6	0.52
25	Reference and User Services Quarterly	17	U S	6	0.52
26	Journal of the Medical Library Association	17	U S	6	0.52
27	Information Today	17	U S	6	0.52
28	Cataloging and Classification Quarterly	17	U S	6	0.52
29	Science	18	U S	5	0.43
30	Nature	18	GB	5	0.43
31	Library Quarterly	18	U S	5	0.43
32	Journal of Experimental Education	18	U S	5	0.43

Cont...

33	Journal of Electronic Publishing	18	U S	5	0.43
34	Issues in Science and Technology Librarianship	18	U S	5	0.43
35	Information Research	18	Sweden	5	0.43
36	IEEE Computer	18	U S	5	0.43
37	Educational Policy Analysis Archives	18	U S	5	0.43
38	Archivaria	18	Canada	5	0.43
39	ACM Computing Surveys	18	U S	5	0.43
40	The Electronic Library	19	GB	4	0.34
41	Science & Technology Libraries	19	U S	4	0.34
42	Machine Translation	19	Netherland	4	0.34
43	Libri	19	Germany	4	0.34
44	Library Hi Tech News, 18	19	GB	4	0.34
45	Journal of the Society of Archivists	19	GB	4	0.34
46	Interactions	19	U S	4	0.34
47	IEEE Trans on information theory	19	U S	4	0.34
48	Cultivate Interactive	19	GB	4	0.34
49	Computer Networks and ISDN Systems	19	U S	4	0.34
50	Title with 1-3 (248)			353	30.30
	Total			**1165**	

Analysis of the cited journal articles was done to determine the productivity of cited authors. A total 2460 authors contributed 1165 cited articles with an average of 2.11 authors per cited articles. Table 8 shows the distribution of the number of cited articles published by each one of the authors and reveals that whereas one author (Van de Sompel, H.) contributed as many as 60 cited articles, others contributed a few.

It can be observed from the table 44.4% of authors contributed to only one article each, those contributing to more than one therefore being much fewer in number. 7.8% and 2.9% authors have contributed two and three articles respectively. Only 15 authors published 10 or more cited articles. The contribution of 24 most prolific authors contributed up to 10 cited articles is shown in table 9.

Table 8 Author Productivity

Number of article	Number of authors	%
1	1092	44.4
2	192	7.8
3	71	2.9
4	19	0.8
5	16	0.6
6	10	0.4
7	7	0.3
8	5	0.2
9	4	0.16
10	4	0.16
11	5	0.2
12	1	0.04
14	3	0.12
16	2	0.08
15	1	0.04
16	1	0.04
17	2	0.08
19	1	0.04
20	2	0.08
21	1	0.04
25	1	0.04
32	1	0.04
35	1	0.04
60	1	0.04

Table 9 Prolific Authors

S. No.	Name of Author	No. of Articles
1	Van de Sompel, H.	60
2	Lagoze, Carl.	35
3	King, Donald W.	32
4	Nelson, M.L.	25
5	Tenopir, C.	21
6	Bekaert, J.	20
7	Montgomery, Carol Hansen	20
8	Arms, W Y	19
9	Payette, Sandra	17
10	Hochstenbach, Patrick	16
11	Liu, X.	16
12	Beit-Arie, O.	14
13	Lynch, Clifford A.	14
14	Smith, M.	14
15	Blanchi, C.	12
16	Hammond, Tony	11
17	Harnard, S.	11
18	Lund, B.	11
19	Warner, S.	11
20	Zia, Lee L.	11
21	Balakireva, L.	10
22	Hannay, Timo	10
23	Maly, Kurt	10
24	Tansley, R.	10

Summary of Findings

The study shows the changing trends of research in the field of library & information science particularly in the field of digital libraries. D-Lib Magazine was found important source of digital library literature. The quality of D-Lib Magazine can be judged on the basis of citations. This magazine cited very good journals not only from library & information science but some related disciplines like computer science. The study shows that 18 countries actively involved in digital library research. With the introduction of Internet and World Wide Web, the trend of scholarly communication has changed. This change can be seen in digital library research as researchers have used digital and web resources to conduct their research.

Average number of citations in the journal is around 16.18 per article. The highest number of citations came during 2002 and the lowest number of citations in the year 2004.

The major sources of information cited in the D-Lib Magazine are web resources and journals articles. Current documents are most frequently cited by the authors of D-Lib Magazine. More than 81 percent literature cited was published during 1996-2006.

It is evidenced that the researchers are mainly consulted journals published from USA and UK. It was observed that 18 countries generated research material on the subject. The countries have been ranked on the basis of frequency of occurrence. It was observed that more than 90 percent of the total cited articles were published from USA and UK. HTML file was found the most frequently used format followed by PDF file format. It was observed that the highest number of domains of the cited document is of .org type, followed by edu/.ac type. www.loc.gov ranks 1st with highest in the rank list of first 37 most cited URLs followed by www.w3.org. The highest number of cited articles with different types of authorship pattern in D-Lib Magazine is 219 in the year 2003, followed by 2002 with 209 articles. Majority of the documents cited have been contributed by single authors. It is important to note that D-Lib Magazine is the most cited source securing 1st position in the cited journals articles and Van de Sompel, H is the most cited authors cited 46 times in D-Lib Magazine, followed by Lagoze, Carl Who was cited 25 times by the authors of D-Lib Magazine.

Suggestions for Further Research
This study was done with D-Lib Magazine which is an open access journal. An exhaustive study is needed to be done to know the research trends in the field of digital library. It is suggested that a comparative study of selected printed journals and online journals in the field of digital library may be conducted to know the trends of research particularly citation impact of printed and online journals in this field.

References

1. Beker, D. R. (1990), "Citation analysis: A methodological review", *Social Work Research & Abstract*,Vol. 26, pp. 3-4.
2. Mukherjee, B. (2009), "The hyperlinking pattern of open-access journals in library and information science: A cited citing reference study", *Library and information science research*,Vol 31, pp. 2-113.
3. Directory of Open Access Journals (DOAJ), available at: http://www.doaj.org/.
4. Open J-gate, available at: http://www.openj-gate.com

About the Authors

Moin Ahmad, received Master of Library and Information Science degree from Aligarh Muslim University, Aligarh and M.Phil degree from Annamali University. He has worked as research intern in the Education, Training and Translation Group, National Institute of Science Communication and Information Resources (NISCAIR), New Delhi and also worked as Assistant Librarian in the Central Library, Integral University, Lucknow. Currently he is working as Librarian in American University of Ras Al Khaimah, UAE.

Mohammad Nazim, is currently working Assistant Librarian in Faculty of Law, Banaras Hindu University, Varanasi. He received Master of Library and Information Science degrees from Aligarh Muslim University, Aligarh and pursuing PhD from the Department of Library & Information Science, Banaras Hindu University. He has published several research papers in reputed national and international journals. His research interests include information retrieval, digital library, e-learning, open access to scholarly communication and knowledge management.