



## A Scientometric Analysis on Global Digital Library Research during 1991-2022

Indrajit Mondal<sup>a</sup> and Shabnam Surendran<sup>b</sup>

<sup>a</sup> Project Engineer, National Digital Library of India, IIT Kharagpur, West Bengal, 721302, India,

<sup>b</sup> Project Engineer, National Digital Library of India, IIT Kharagpur, West Bengal, 721302, India,

Email: surendranshabnam@gmail.com; mondal.indrajit.709406@gmail.com

Received: 21 February 2024; Accepted: 09 April 2024

This study conducts a comprehensive scientometric analysis of global digital library research from 1991 to 2022. The research employs the BibliometrixR and Vosviewer packages to analyze a diverse set of 6157 scholarly publications from the Scopus database and evaluate key bibliometric indicators i.e., publication output, collaboration networks, citation patterns, and thematic trends. The findings show that the field of digital library research experienced exponential growth, with the highest number of publications in 2022 (9.53%). Lecture Notes in Computer Science (12.81%), Proceedings of the ACM/IEEE Joint Conference on Digital Libraries (5.77%), and D-Lib Magazine (3.43%) are the top-ranked journals during this time. The top prolific authors in digital library research are Nelson ML from Old Dominion University (USA) with 54 publications and Fox EA from Virginia Polytechnic (USA) with 50 publications. The top most relevant affiliations are the Nanyang Technological University of Singapore with 191 articles and the University of California with 172 articles. The USA ranks first in scientific production, followed by the UK and India. The top most cited countries are the USA with 26535 citations and India with 635 citations. Keyword co-occurrence analysis reveals themes in digital library research, including digital libraries, information retrieval, systematic review, and artificial intelligence.

**Keywords:** Digital Library; Electronic Library; Bradford's Law; Scientometrics; BibliometrixR; Vosviewer

### Introduction:

The aim of scientific communication is to share research findings with the entire scientific community and its peers. One of the essential elements in quantifying research output in the academic environment is research publications. It provides knowledge to young researchers to disseminate and reuse for societal advancement. Scientometric studies in a research field reflect the growth of the literature in the field. Using indicators like publication counts, citation counts, keyword analysis, co-citation analysis, and other metrics, one can analyze the research contribution in a field. According to Subramanyam (Subramanyam, 1983), the assessment of collaboration ought to emphasize a holistic approach. He also claimed that the complexity of human interaction makes it challenging to quantify the nature of research collaboration.

Digital libraries have transformed information dissemination and accessibility, becoming pivotal in the modern information age. The significance of digital library research lies in its ability to enhance access, preservation, and management of digital information resources.

### Literature Review:

Digital libraries are a rapidly growing area of research in the fields of librarianship and information science, and some scientometric studies have recently been done.

Sutradhar *et al.*'s (2022) study on Immersive Library Environment to Access Virtual Reality Content shows that the Central Library of IIT Kharagpur uses web-based augmented reality, virtual reality, and data visualization services. A survey reveals 90% of users are satisfied, with data privacy, health, and safety being critical factors.

Jalal and Mukhopadhyay (2022) analyzed large-scale data cleaning and analysis of mathematics publication data from Scopus, revealing a significant gender disparity in the field, with 73% male authors and 27% female authors, highlighting the need for improved data management and analysis.

Jalal (2019) analyzed 1156 papers from 1991 to 2017 between India and Bangladesh, revealing major collaborations in medical, agriculture, and biological sciences. The research network was built using BibliometrixR and VOSviewer, and the study confirmed the validity of Lotkas's Law for international co-authorship data.

Shukla and Verma (2019) conducted a scientometric assessment of digital library research in India from 1989 to 2018 using Scopus and discovered that the highest 108 (10.11%) of research papers were published in the year 2016, and the maximum 398 research papers were published by two authors. Shalini Urs was the most prolific author, with 13 publications.

A study by Ahmed *et al.* (2018) revealed that from 2002 to 2016, 4,206 digital library-related literatures were published worldwide, with the USA contributing the most at 38.94%. The most prolific authors were from the USA and the UK, Brazil, South Africa, China, and Germany, with three each from the USA and the UK.

Sood *et al.* (2018) analyzed digital library publications from 2006 to 2015, finding the highest number of papers in 2013, followed by 2007 with Fox E A and Urs S R producing the most.

Gupta *et al.* (2017) analyzed global digital library research output from 2007-16, revealing the U.S. contributed the most with 26.89%, while the computer science discipline had the highest contribution at 30.86%.

Singh *et al.*'s (2007) study on research productivity in digital libraries from 1998-2004, using the LISA database, found that 61% of articles were single-authored and did not align with Lotka's Law. The majority of articles were found in the D-Lib Magazine journal.

Cunningham (1998) study explores the effectiveness of digital libraries in bibliometric analysis and their limitations. It highlights the potential for new studies, including full text analysis, access to technical reports or pre-prints, and direct measurement of document usage.

The above review of the literature reveals various scientific studies conducted by authors to analyze authorship patterns, collaborative measures, and article influence over the year, all relevant to the objectives of this study.

### Objectives of the study:

The main objectives of this scientometric analysis in global digital library research output are to identify:

1. Year-wise distribution
2. Top 20 preferred journals
3. Top 20 prolific authors
4. Top 20 most relevant affiliations
5. Top 20 country's scientific production

6. Top 25 country's citation patterns
7. Collaboration networks
8. Thematic trends

### Methodology:

The study aims to elucidate the trends, strengths, and thematic orientations of digital library research around the world. For this, research done on digital libraries from "all regions" is considered. The data has been downloaded from the "Scopus database" on 17<sup>th</sup> August 2023. To retrieve the dataset for conducting the current study, the following search strategy is used: TITLE-ABS-KEY= ("Digital Library" OR "Electronic Library" OR "Virtual Library") further refined by the time span "1991-2022", and language as "English". Only articles, reviews, proceedings papers, and book chapters have been included. Document types such as editorials, notes, and corrections were excluded from the study. A total of 6157 records that were retrieved using the search query have been analyzed.

The scientometric analysis has been conducted using the BibliometrixR (Biblioshiny) and Vosviewer packages. This encompassed bibliometric indicators such as publication volume, collaboration patterns, citation metrics, and the co-occurrence of keywords.

### Results

#### 1. Year-wise distribution of publication:

The year-wise distribution of the publications in Table 1 shows exponential growth in this field with the highest number of publications in the year 2022 (9.53% of total publications) followed by years 2021 (8.19%) and 2020 (7.73%). Figure 1 shows the corresponding graphical representation.

#### 2. Top 20 preferred journals:

In the list of top 20 journals in Table 2a based on number of publications indexed in the Scopus database during 1991-2022, it is found that 'Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)' has published 789 (12.81%) of total publications, followed by 'Proceedings of the ACM/IEEE Joint Conference on Digital Libraries' with 355 (5.77%) publications and 'D-Lib Magazine' with 211 (3.43%) publications. Figure 2 satisfies Bradford's law regarding the contribution of core journals to the publication outputs of this field. Table 2b shows the

Table 1 — Year-wise distribution of publications.

Year	Articles	Percentage of Publication	Cumulative	Year	Articles	Percentage of Publication	Cumulative
1991	4	0.07	0.07	2007	253	4.11	26.22
1992	1	0.02	0.09	2008	178	2.89	29.11
1993	2	0.03	0.12	2009	214	3.48	32.59
1994	4	0.07	0.19	2010	172	2.79	35.38
1995	28	0.46	0.65	2011	195	3.17	38.55
1996	30	0.49	1.14	2012	200	3.25	41.8
1997	38	0.62	1.76	2013	192	3.12	44.92
1998	56	0.91	2.67	2014	222	3.61	48.53
1999	72	1.17	3.84	2015	247	4.01	52.54
2000	82	1.33	5.17	2016	270	4.39	56.93
2001	98	1.59	6.76	2017	296	4.81	61.74
2002	128	2.08	8.84	2018	363	5.90	97.64
2003	162	2.63	11.47	2019	428	6.95	74.59
2004	148	2.40	13.87	2020	476	7.73	82.32
2005	237	3.85	17.72	2021	504	8.19	91.51
2006	270	4.39	22.11	2022	587	9.53	~100

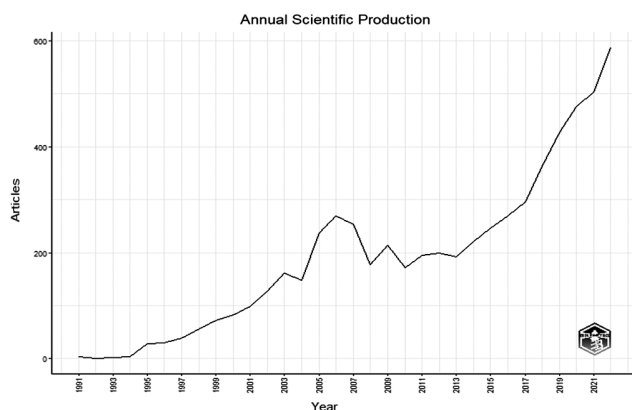


Fig. 1 — Annual scientific production.

zone-wise summary for total number of journals and publications.

### 3. Top 20 prolific authors:

In Table 3, among the top 20 most prolific authors in digital library research, it is found that '*Nelson ML*' from Old Dominion University (USA) has occupied the *1st* rank with 54 publications, '*Fox EA*' from Virginia Polytechnic (USA) occupied the *2nd* rank with 50 publications.

### 4. Top 20 most relevant affiliations:

In Table 4, within the top 20 most relevant affiliations, '*Nanyang Technological University*' of 'Singapore' occupied the *1st* rank with 191 articles, followed by '*University of California (USA)*' with 172 articles.

### 5. Top 20 Country's Scientific Production:

Table 5 shows the country's scientific production. '*USA*' has occupied the *1st* rank, followed by '*UK*' with the *2nd* rank and '*India*' with the *12th* rank.

### 6. Top 25 Country's Citation Patterns:

In Table 6, among the top 25 most cited countries list, '*USA*' has occupied the *1st* rank with 26535 citations, and '*India*' has occupied the *21st* rank with 635 citations.

### 7. Collaboration Networks:

It is a prominent feature of digital library research among all countries around the world. Researchers frequently engage in international collaborations, fostering cross-border knowledge exchange and enriching the research landscape. Figure 3 shows the co-authorship network among the most contributing countries. Here it shows the '*USA*' has the maximum co-authorship network, followed by the '*UK*' and '*Brazil*'.

### 8. Thematic Trends:

In thematic trends, keyword co-occurrence analysis in Figure 4 reveals the themes within digital library research with the most frequent word '*digital libraries*' followed by '*information retrieval*', '*systematic review*' and '*artificial intelligence*'.

The word cloud in Table 7 shows the thematic diversity and evolving research priorities within the field.

Table 2a — Top 20 preferred journals.

Sources	Rank	Freq.	Cum. Freq.	Zone
Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	1	789	789	Zone 1
Proceedings of the ACM/IEEE Joint Conference on Digital Libraries	2	355	1144	Zone 1
D-Lib Magazine	3	211	1355	Zone 1
Proceedings of the ACM International Conference on Digital Libraries	4	104	1459	Zone 1
ACM International Conference Proceeding Series	5	102	1561	Zone 1
Communications in Computer and Information Science	6	93	1654	Zone 1
Journal of Physics: Conference Series	7	76	1730	Zone 1
IEEE Access	8	66	1796	Zone 1
International Journal on Digital Libraries	9	55	1851	Zone 1
Proceedings of the Asist Annual Meeting	10	55	1906	Zone 1
Journal of Medical Internet Research	11	48	1954	Zone 1
Communications of the ACM	12	47	2001	Zone 1
Procedia Computer Science	13	47	2048	Zone 1
Library Review	14	40	2088	Zone 2
Journal of Chemical Information and Modeling	15	39	2127	Zone 2
Library Hi Tech	16	36	2163	Zone 2
BMJ Open	17	35	2198	Zone 2
Information Technology and Libraries	18	35	2233	Zone 2
Liber Quarterly	19	34	2267	Zone 2
Information Services and Use	20	32	2299	Zone 2
Proceedings of the Association for Information Science And Technology	21	31	2330	Zone 2
Electronic Library	22	30	2360	Zone 2
IFIP Advances in Information and Communication Technology	23	26	2386	Zone 2
Plos One	24	26	2412	Zone 2
Journal of the American Society for Information Science and Technology	25	25	2437	Zone 2

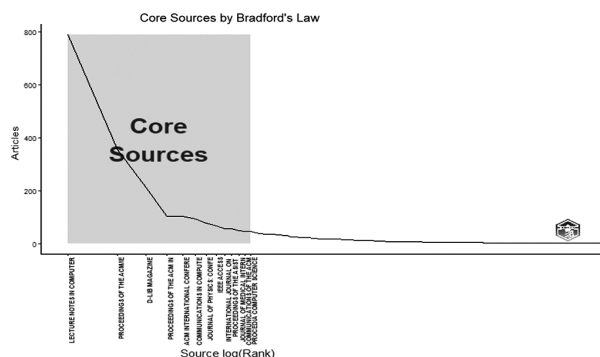


Fig. 2 — Contribution of core sources and Bradford's law.

**Discussion:**

The year-wise distribution of publications shows exponential growth in this field, with the highest number of publications in the year 2022 (9.53% of total publications).

In the list of top 20 journals based on the number of publications indexed in the Scopus database during 1991-2022, it is found that 'Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)' has published 789 (12.81%) of total publications, followed by 'Proceedings of the

Table 2b — Zone-wise summary table for total number of journals and publications.

Zone	Number of journals	Number of publications
1	13	2048
2	228	2048
3	1576	2061
Total = 1817		Total = 6157

ACM/IEEE Joint Conference on Digital Libraries' with 355 (5.77%) publications and 'D-Lib Magazine' with 211 (3.43%) publications.

In the top 20 most prolific authors in digital library research, it is found that 'Nelson ML' from Old Dominion University (USA) has occupied the 1st rank with 54 publications, and 'Fox EA' from Virginia Polytechnic (USA) has occupied the 2nd rank with 50 publications.

Among the top 20 most relevant affiliations, 'Nanyang Technological University' of 'Singapore' occupied the 1st rank with 191 articles, followed by 'University of California (USA)' with 172 articles.

In the country's scientific production, 'USA' has occupied the 1st rank, followed by 'UK' with the 2nd rank and 'India' with the 12th rank.

Table 3 — Top 20 prolific authors.

Sl. No.	Authors	Affiliation	Articles
1	Nelson ML	Old Dominion University, USA	54
2	Fox EA	Virginia Polytechnic, USA	50
3	Bainbridge D	University of Waikato, New Zealand	49
4	Witten IH	University of Waikato, New Zealand	43
5	Lim E-P	Singapore Management University, Singapore	34
6	Cunningham SJ	University of Waikato, New Zealand	33
7	Goh DH-L	Nanyang Technological University, Singapore	33
8	Hinze A	University of Waikato, New Zealand	33
9	Giles CL	Pennsylvania State University, USA	31
10	Mayr P	GESIS Leibniz Institute for Social Sciences, Germany	31
11	Buchanan G	University of Melbourne, Australia	29
12	Gipp B	University of Gottingen, Germany	29
13	Blandford A	University College London, UK	27
14	Van De Sompel H	DANS & Ghent University, Belgium	27
15	Zhang Y	Rutgers University, USA	24
16	Ferro N	University of Padua, Italy	22
17	Nichols DM	University of Waikato, New Zealand	22
18	Theng Y-L	Nanyang Technological University, Singapore	22
19	Klein M	Los Alamos National Laboratory, USA	21
20	Rauber A	Vienna University of Technology, Austria	21

Table 4 — Top 20 most relevant affiliations

Sl. No.	Affiliation	Country/Region	Articles
1	Nanyang Technological University	Singapore	191
2	University of California	USA	172
3	University of Waikato	New Zealand	172
4	University of Maryland	USA	83
5	University College London	UK	79
6	Los Alamos National Laboratory	USA	71
7	Old Dominion University	USA	70
8	University of Padua	Italy	61
9	University of Southampton	UK	61
10	Pennsylvania State University	USA	60
11	Universidade De Sao Paulo	Brazil	58
12	Indiana University	USA	56
13	Carnegie Mellon University	USA	54
14	National University of Singapore	Singapore	54
15	University of Strathclyde	UK	54
16	University of North Carolina	USA	53
17	Cornell University	USA	49
18	University of Glasgow	UK	47
19	University of Pittsburgh	USA	47
20	Texas A and M University	USA	46

Table 5 — Top 20 countries' scientific production.

Sl. No.	Region	Frequency	Sl. No.	Region	Frequency
1	USA	5796	11	Australia	493
2	UK	2011	12	India	464
3	Brazil	1900	13	Singapore	388

(contd.)

Table 5 — Top 20 countries’ scientific production. (contd.)

Sl. No.	Sl. No.	Sl. No.	Sl. No.	Sl. No.	Sl. No.
4	Germany	1375	14	New Zealand	360
5	China	1176	15	Japan	347
6	Italy	878	16	Switzerland	320
7	France	722	17	Greece	286
8	Spain	697	18	Portugal	275
9	Canada	591	19	Saudi Arabia	241
10	Netherlands	514	20	Malaysia	240

Table 6 — Top 25 country’s citation patterns.

Sl. No.	Country	TC	Average Article Citations	Sl. No.	Country	TC	Average Article Citations
1	USA	26535	24.6	14	Singapore	1280	17.5
2	United Kingdom	10379	26.2	15	Greece	970	17.3
3	Germany	7154	30.1	16	New Zealand	857	12.2
4	Brazil	3656	11.4	17	Japan	829	15.9
5	Canada	2919	25.4	18	Ireland	798	16
6	France	2635	21.8	19	Saudi Arabia	748	17.8
7	China	2535	9.7	20	Portugal	648	12.7
8	Australia	1872	20.3	21	India	635	10.4
9	Italy	1853	10.2	22	Malaysia	627	13.9
10	Norway	1848	71.1	23	Hong Kong	522	52.2
11	Spain	1802	17.8	24	Sweden	493	16.4
12	Netherlands	1739	21.2	25	South Korea	451	12.5
13	Switzerland	1458	27.5				

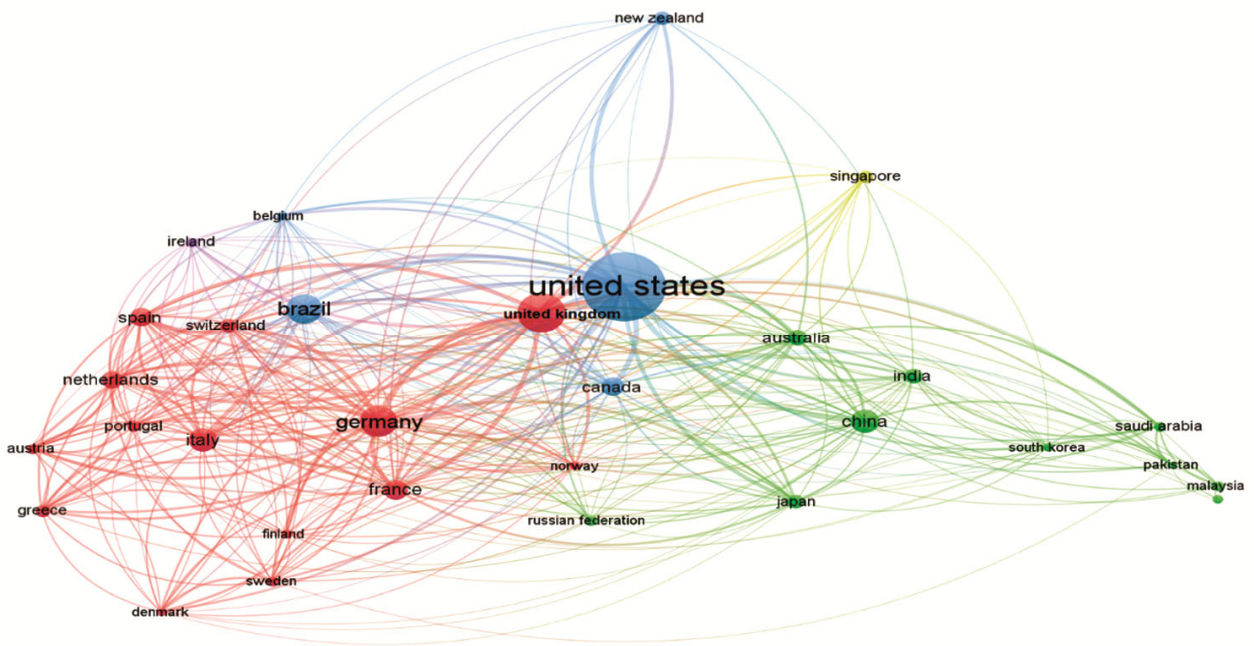


Fig. 3 — Country’s co-authorship network.



contributions, collaboration networks, citation impact, and thematic foci. This study contributes to a deeper understanding of the dynamics within digital library research, guiding future research endeavours and strategic initiatives.

#### Acknowledgements:

We would like to express our sincere gratitude to Ms. Suparna Chaudhuri, Technical Officer, National Digital Library of India, IIT Kharagpur for her valuable guidance and support.

#### Declarations:

◆ **Conflict of interest:** The authors declare no relevant conflicts of interest.

#### References:

- 1 Roy S G, Kanjilal U, Sutradhar B and Jalal S K, Building Immersive Library Environment to Access Virtual Reality Content - A Proposed Framework Model, *DESIDOC Journal of Library & Information Technology*, 42 (3) (2022)178-184.
- 2 Jalal S K, and Mukhopadhyay P, Gender differences, data carpentry and bibliometric studies in Mathematics, *COLLNET Journal of Scientometrics and Information Management*, 16 (2) (2022)465-476.
- 3 Jalal S K, Co-authorship and co-occurrences analysis using BibliometrixR package: a case study of India and Bangladesh, *Annals of Library and Information Studies*, 66 (June) (2019) 57-64.
- 4 Shukla R and Verma M K, Digital library research in India during 1989-2018: a scientometric analysis based on Scopus database, *Journal of Information & Systems Management*, 9 (2) (2019)62-73.
- 5 Ahmad K, Ming Z J and Rafi M, Assessing the digital library research output: Bibliometric analysis from 2002 to 2016, *Electronic Library*, 36 (4) (2018)696-704.
- 6 Sood S, Tiwari D and Khanna S, Research in digital library field: a scientometric assessment of publications output during 2006-15, *International Journal of Library Information Network and Knowledge*, 3 (2) (2018)76-93.
- 7 Gupta B M, Dhawan S M, Gupta R and Faizul N, Digital libraries: A scientometric assessment of global publications output during 2007-16, *International Journal of Information Dissemination and Technology*, 7 (4) (2017)247-252.
- 8 Singh G, Mittal R and Ahmad M, A bibliometric study of literature on digital libraries. *The Electronic Library*, 25 (3) (2007)342-348.
- 9 Subramanyam K, Bibliometric studies of research collaboration: A review, *Journal of Information Science*, 6 (1) (1983)33-38.
- 10 Cunningham S J, Applications for bibliometric research in the emerging digital libraries, *Scientometrics*, 43 (2) (1998) 161-175.