

Global Retracted Research Publications: A Bibliometric Study

Mohit Garg and Nabi Hasan

The present study used a bibliometric technique to draw the patterns in retracted publications over time. A sample of bibliographic data of 12876 documents up to the year 2021 was extracted from the Scopus database. The analysis found that the number of retracted documents has significantly increased in the last decade, with most documents in 2010 (4377) followed by 2011 (3534). The study found that most retracted documents were from China, followed by the United States and India. The highest number of 70 retracted documents were from the author Nazari, A. of Islamic Azad University, Iran, with a rate of 4.04 retractions per total document published. The other authors were Ueshima, H. (68), Otake, H. (45), Sato, Y. (45), and Iwamoto, J. (41) from Japanese Institutions. The highest rate of retraction per total documents were of Ghoranneviss, M., Islamic Azad University, Iran (24.65), followed by Boldt, J., Justus-Liebig-Universität Gießen (24.04), and Shamshirband, S. National Yunlin University of Science and Technology, Taiwan (18.54). The majority of the source type was found to be Conference Proceedings (11096, 60.71%), followed by Journal Publications (7159, 39.17%). The study found the maximum number of retracted publications were in the Conference Proceedings (11096, 60.71%), followed by Journals (7159, 39.17%).

Introduction

The scientific community conducts rigorous work in their laboratory to create new knowledge. This knowledge solves or provides the optimal solution to the existing problem. The scholars disseminate the outcomes of their study to the respective peer group as the scientific publications in Journals, Conferences, Books, etc. This research relies on various ethical parameters and the integrity of the scholar. However, some scholars ignore these, do different types of fraud, and violate the ethics of research and publications. Nowadays, these are the concerns of the global scientific community. These issues are reported by the scientific community or sometimes self-reported by the authors [1]. Retraction is a practice of ethical misconduct in scientific communication. It can be occurred due to Data Falsification or manipulation, ethical misconduct like plagiarism, errors in data or methods

2. Literature Review

A large number of systematic and bibliometric reviews have been published. A few of such related studies are highlighted in this paper.

Samp et.al [2] investigated the scientific publications on drug literature from 2000 to 2011. They extracted the data from the PubMed database. They found that 102 articles were related to the drug or biomedical field.

The study found that majority of the publications (73, 72%) were retracted due to scientific misconduct, while only 29 studies were retracted due to error.

Rosenkrantz [3] used the PubMed database to identify the trends in retracted publications in radiology journals. The author found that 1.1% of PubMed publications have received retraction notices. The study found the key reasons for retraction were improper methods or results, duplication of publications, plagiarism, permission issue, publisher error, etc.

Rai & Sabharwal [4] assessed the 59 retracted publication to discover the trends and characteristics of scientific productivity in the field of orthopedics. These articles were traced from five databases (PubMed, Google Scholar, CINAHL, Scopus, and MEDLINE). These databases were searched with the query “Retracted Publication Orthopedic.”. They identified that three articles per 10,000 were retracted/

Wang et.al [5] examined retracted documents of the biomedical domain published in Open Access Journals. The authors have extracted the data of 621 publications from the Medline database. They found that number of retractions had been increased in the Open Access Journals. They identified that most of the retracted publications were from researchers of China, India, Iran, and the USA published in low-impact factor journals.

Chauvin et.al [6] investigated 28 retracted notes of the publications in the field of emergency medicine indexed in MEDLINE, Web of Science, and Cochrane Central Register of Controlled Trials. The study found that retractions in emergency medicine were less in number. The authors identified that majority of the notices were due to plagiarism followed by duplicate publication and overlap.

Chambers et.al [7] examined the reasons for retractions in the domain of obstetrics and gynecology. They have extracted the data of retracted articles from the PubMed database. They have identified 176 retracted articles in two years time frame. They found that only 18 authors had two or more retractions. The authors found that the major reasons for retraction were Plagiarism and data falsification.

Samuel et.al [8] investigated 8092 articles indexed in the PubMed database in the field of dentistry. The authors have used panda’s library in python for the analysis. They found that the maximum retraction of articles in dentistry-related research originated from India (25.3%). They also found that the average time for the issue of retraction notice was 2.6 years. They identified plagiarism as the primary reason for retraction notices. The non-funded research (62.16%) published in low-impact factor journals have the most number of retracted documents.

Zhang & Fu [9] analyzed 1839 retracted articles in Clinical Medicine. The authors have extracted the data from the Sci-Expanded Web of Science database from 1978 to 2020. They have identified the six main reasons for retraction. These are Falsification, Errors, Self-Plagiarism, Ethical Issues, Plagiarism, and Authorship Issues. They found a close collaboration between the authors with multiple retracted publications.

Brown et.al [10] conducted systematic reviews of the 1396 retracted publications of the pharmacy domain indexed in the Web of Science. They found that most of the citations of the retracted publication were received before the retraction.

The above review shows that many domain-specific studies have been conducted on the analyses of retracted publications. Thus, there is a need for a study that explores the pattern in the worldwide publication of retracted work. Therefore, we conducted a comprehensive bibliometric analysis of global retracted publications in this study. The present study aimed to identify different patterns in the context of global retracted publications.

3. Objective

To explore the evolution of global retraction publication.

To identify the author/institutions/countries that published most retracted publications.

To identify the sources published in most retracted publications

To portray the subject and themes of retracted work

4. Methodology

In this study, we have used bibliometric techniques to portray the patterns in global retracted publications. The Scopus was considered a data source for the present study. Scopus is a popular bibliographic citation database product by Elsevier. It was considered over other databases like the web of science, as it has more coverage than others. This phenomenon has been studied by different scholars [11,12]

The bibliographic data of documents that have been marked under the category of “retracted” was extracted from the Scopus database on 1 June 2022. The Scopus database defines 13 types of document types indexed in it. These are Article, Article-in-press (AiP), Book, Chapter, Conference Paper, Data Paper, Editorial, Erratum, Letter, Note, Retracted article, Review, and Short Survey. The retracted article is one such document type defined by Scopus and represented by the word “tb”. The Scopus defined retracted article as “Published articles that the author(s) or publisher has requested to retract” [13].

The data was limited to the timeframe parameters up to 2021, and the document’s language was English. The final query for the search on Scopus was (DOCTYPE(tb) AND (EXCLUDE (PUBYEAR,2022)) AND (LIMIT-TO(LANGUAGE, “English”)))

The data includes information of authors, affiliations, country, title, source, citation, etc., and was exported in CSV format. Microsoft Excel performed the analysis of the year-wise distribution of documents, Authors/Countries/Institutions with most retraction, Subject Area, Publication Venues, etc., in tabular format. The open-source environment R was used for plotting the temporal data and a word cloud of the keyword.

5. Results

The bibliographic data of a total of 18276 documents were extracted. Figure 1 shows the year-wise trends of growth of retracted documents presented in the form of a line plot with three different levels, i.e., Red, Green, and Blue. The first level is represented by Red Color with retracted documents less than or equal to 39 per year. This was the timeframe when a minimal number of retracted documents were noticed. The evolution of retracted documents The first retracted documents appeared in 1975, with the title “Stickler syndrome Report of a second Australian family,” published in Pediatric Radiology.

After 14 years, In the year 1989, two documents were retracted. In fact, after this, each year has observed some documents lensed under retraction. The year 2005 observed the first steep rise in the retraction of documents. After this, there has been continuous exponential growth in the retracted documents. The green color zone with less than or equal to 964 and more than 39 retracted documents is represented by the years 2005 to 2009, 2012 to 2014, 2016, 2020 to 2021. The highest number of retracted documents was observed in 2010 (4377,23.95%), followed by 2011(3534,19.34%). The analysis found a sudden decline in the retraction of documents in 2012, with just one-tenth of the previous year’s documents (387,2.12%). From 2012 to 2021, the next nine years observed less retraction, but it has remained close to 964 documents. The year-wise distribution of the retracted documents was also presented in tabular format, as shown in table 1.

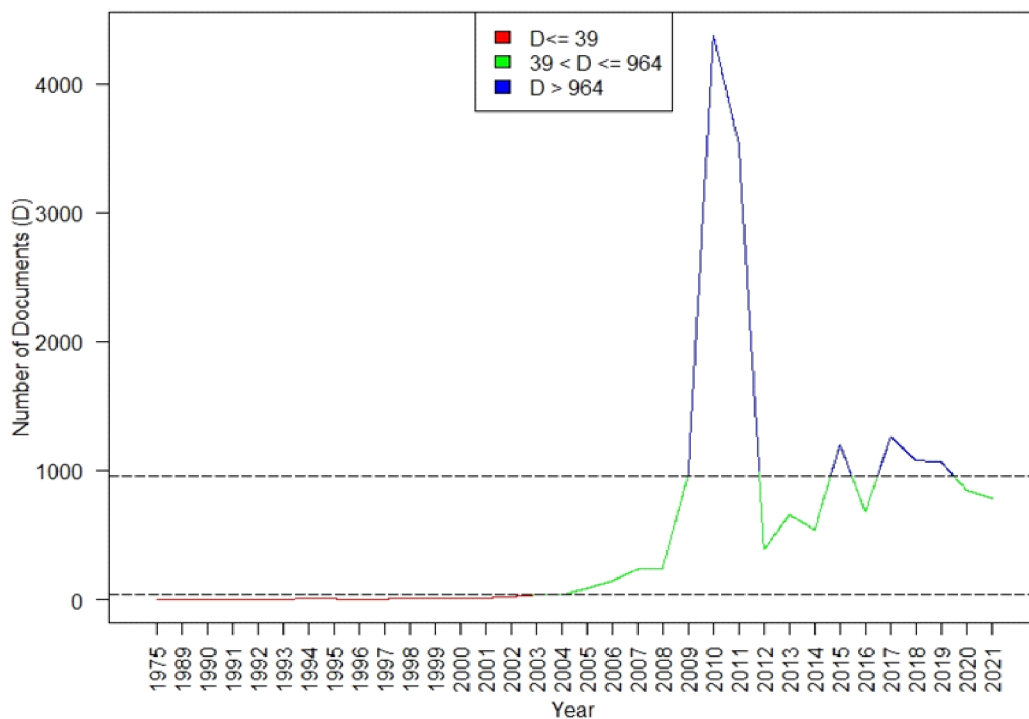


Figure 1: Annual Distribution of Retracted Documents

Table 1: Year-wise distribution of retracted documents

Year	Number of Documents	%	Cumulative %
1975	1	0.01	0.01
1989	2	0.01	0.02
1990	1	0.01	0.02
1991	1	0.01	0.03
1992	1	0.01	0.03
1993	6	0.03	0.07
1994	7	0.04	0.10
1995	7	0.04	0.14
1996	3	0.02	0.16
1997	6	0.03	0.19
1998	9	0.05	0.24
1999	7	0.04	0.28
2000	11	0.06	0.34
2001	14	0.08	0.42
2002	23	0.13	0.54
2003	39	0.21	0.76
2004	37	0.20	0.96
2005	86	0.47	1.43
2006	145	0.79	2.22
2007	237	1.30	3.52
2008	244	1.34	4.85
2009	964	5.27	10.13
2010	4377	23.95	34.08
2011	3534	19.34	53.41
2012	387	2.12	55.53
2013	661	3.62	59.15
2014	542	2.97	62.11
2015	1204	6.59	68.70
2016	678	3.71	72.41
2017	1263	6.91	79.32
2018	1079	5.90	85.23
2019	1067	5.84	91.06
2020	848	4.64	95.70
2021	785	4.30	100.00

Table 2: Authors with most Retracted Documents

The authors were ranked according to the number of documents that have received retraction.

Sl. No.	Author Name	Number of retracted documents	H Index	Total number of documents in Scopus	Citations	Retracted paper per total number of documents	Affiliation	Country
1	Nazari, A.	70	41	283	6186	4.04	Islamic Azad University	Iran
2	Ueshima, H.	68	21	258	1484	3.79	Showa University Hospital	Japan
3	Otake, H.	45	19	207	1098	4.60	Showa University School of Medicine	Japan
4	Sato, Y.	45	45	246	6025	5.47	Mitate Hospital	Japan
5	Iwamoto, J.	41	42	285	6025	6.95	Keiyu Orthopaedic Hospital	Japan
6	Sarkar, F.H.	37	109	618	41606	16.70	Wayne State University School of Medicine	US
7	Maxim, A.	36	11	83	413	2.31	Silicon Laboratories,	US
8	Riahi, S.	36	29	99	2634	2.75	Islamic azad University	Iran
9	Orlandi, A.	33	31	337	4928	10.21	Università degli Studi dell'Aquila	Italy
10	Banerjee, S.	30	72	162	15700	5.40	Wayne State University	US
11	Salar Elahi, A.	28	22	232	1944	8.29	Islamic Azad University	Iran
12	Wang, Z.	26	78	239	18019	9.19	Bengbu Medical College	China
13	Boldt, J.	25	52	601	10555	24.04	Justus-Liebig-Universität Gießen	Germany
14	Fujii, Y.	24	29	340	3117	14.17	Faculty of Medicine, University of Tsukuba	Japan
15	Shamshirband, S.	24	59	445	13033	18.54	National Yunlin University of Science and Technology	Taiwan
16	Ghoranneviss, M.	23	31	567	4976	24.65	Islamic Azad University	Iran
17	Petkoviæ, D.	23	45	216	6225	9.39	University of Niš	Serbia
18	Calvo-Guirado, J.L.	22	32	202	3457	9.18	Universidad Católica de Murcia	Spain
19	Abdullah, D.	21	1	43	7	2.05	University of Baghdad	Iraq
20	Takeda, T.	21	28	146	3067	6.95	Keio University School of Medicine	Japan

The authors were ranked according to the number of documents that have received retraction. The authors with at least 21 retracted documents are listed in Table 2. Nazari, A from Islamic Azad University, Iran, has received the highest number of retractions with 70 documents, followed by Ueshima, H. of Showa University

Hospital, Japan, with 68 documents. However, the retraction rate per total number of documents for both authors is close to 4 (4.04 & 3.79). The highest rate of retraction per total number of the document was found by Ghoranneviss, M., of Islamic Azad University, Iran (24.65), followed by Boldt, J. of Justus-Liebig-Universität Gießen, Germany (24.04), Shamsheirband, S. of National Yunlin University of Science and Technology, Taiwan (18.54). The majority of the authors were from Japan (06), followed by Iran (04), the United States (03), and one each in China, Germany, Iraq, Italy, Serbia, Spain, and Taiwan.

Table 3: Affiliation-wise retracted documents

Sl. No.	Affiliation	Country	Number of Documents
1	Chinese Academy of Sciences	China	317
2	Ministry of Education China	China	308
3	Wuhan University of Technology	China	212
4	Wuhan University	China	198
5	Tongji University	China	163
6	Huazhong University of Science and Technology	China	159
7	Jilin University	China	157
8	Zhejiang University	China	148
9	Harbin Institute of Technology	China	148
10	Tianjin University	China	130
11	Shanghai Jiao Tong University	China	120
12	Northwestern Polytechnical University	China	115
13	Central South University	China	114
14	Zhengzhou University	China	114
15	North China Electric Power University	China	114
16	Henan Polytechnic University	China	113
17	Shandong University	China	106

Table 3 shows the institution which has more than 100 retracted documents. The analysis found that all the top seventeen institutions were from China. The Chinese Academy of Sciences (317) tops the list with the most retracted documents, followed by the Ministry of Education China (308).

Table 4 presents the top twenty countries regarding the number of retracted documents. As indicated in previous table 3, the analysis found the most number of retractions were from China (11601,63.48%) followed

by the United States (1580, 8.65%) and India (1487, 8.14%), which is approximately one-tenth of retracted documents of China.

Table 4: Country Wise Retracted Documents

Sl. No.	Country	Number of Documents	%
1	China	11601	63.48
2	United States	1580	8.65
3	India	1487	8.14
4	Japan	556	3.04
5	Iran	554	3.03
6	United Kingdom	328	1.79
7	South Korea	308	1.69
8	Italy	262	1.43
9	Taiwan	252	1.38
10	Germany	240	1.31
11	France	236	1.29
12	Malaysia	216	1.18
13	Canada	194	1.06
14	Australia	157	0.86
15	Pakistan	156	0.85
16	Spain	129	0.71
17	Saudi Arabia	122	0.67
18	Egypt	120	0.66
19	Turkey	116	0.63
20	Netherlands	104	0.57

Table 5 shows the subject area with the highest number of retractions. The area of Computer Science and Engineering had received the highest number of retractions, followed by Medicine, Biochemistry, Genetics, and Molecular Biology. The least number of retracted documents were found in Social Sciences, Arts, and Humanities.

The sources with the highest number of retracted documents are presented in Table 6. It was found that the majority of the sources were journals, and only one was Conference Proceedings in the list. The highest

number of retracted documents were published in Plos One (298) followed by the Journal of Physics Conference Series (296). Most of the sources were in the domain of Biochemistry, Genetics, and Molecular Biology, followed by Multidisciplinary. The study found retracted documents in Nature Journal (23), which have a Journal Cite Score of 70.2. It shows that the publication in even the most reputed journals also faces retraction.

Table 5: Subject Area Wise Retracted Documents

Sl. No.	Subject Area	Number of Documents
1	Computer Science	5586
2	Engineering	5258
3	Medicine	3196
4	Biochemistry, Genetics and Molecular Biology	2274
5	Business, Management and Accounting	2147
6	Physics and Astronomy	1982
7	Decision Sciences	1674
8	Health Professions	1180
9	Social Sciences	1077
10	Materials Science	1033
11	Mathematics	969
12	Environmental Science	809
13	Energy	767
14	Agricultural and Biological Sciences	679
15	Chemistry	608
16	Pharmacology, Toxicology and Pharmaceutics	527
17	Multidisciplinary	467
18	Chemical Engineering	458
19	Economics, Econometrics and Finance	448
20	Earth and Planetary Sciences	420
21	Neuroscience	388
22	Immunology and Microbiology	372
23	Psychology	150
24	Arts and Humanities	93
25	Nursing	82
26	Dentistry	55
27	Veterinary	32

Table 7 presents the category of sources published in all retracted documents. The analysis found that the highest number of retracted documents were from Conference Proceedings (11096, 60.71%), followed by Journal (7159, 39.17%) and Book Series (20, 0.11%).

Figure 2 shows the word cloud of the Author's publication keywords. Before plotting the word cloud, similar words are mapped to one common word. For example, humans merged with humans; Animals and animals merged with animals; and so on. The one big word, "Epithelial Mesenchymal Transition" could not be included in the word cloud due to the large number of characters. The big font size of the word represents the highest number of occurrences of the keyword. Figure 2 illustrates that most of the retracted studies focused on human and animal-related research.

Table 6: Popular Sources of Retracted Documents

Sl. No.	Source Title	Number of Documents	Subject Area	2021 Cite Score	2021 SJR	2021 SNIP	Source Type
1	Plos One	298	M	5.6	0.852	1.368	J
2	Journal Of Physics Conference Series	296	PA	0.8	0.21	0.395	CP
3	Journal Of Cellular Biochemistry	189	BGMB	10.1	0.856	0.909	J
4	Arabian Journal Of Geosciences	158	EPS	2.3	0.406	0.831	J
5	Tumor Biology	147	BGMB	6	0.643	0.903	J
6	Scientific Reports	95	M	6.9	1.005	1.389	J
7	Journal Of Cellular Physiology	74	BGMB	13.4	1.308	1.348	J
8	Journal Of Clinical Anesthesia	70	M	6.7	1.081	1.581	J
9	IEEE Transactions On Electromagnetic Compatibility	57	E	4.8	0.854	1.491	J
10	Biomedicine And Pharmacotherapy	41	PTP	11.6	1.194	1.517	J
11	Journal Of Biological Chemistry	41	BGMB	8.8	1.871	1.239	J
12	Neural Computing And Applications	41	CS	8.7	1.072	1.653	J
13	Biomed Research International	39	BGMB	5	0.647	1.119	J
14	Artificial Cells Nanomedicine And Biotechnology	38	PTP	13.6	0.884	1.188	J
15	Cancer Research	34	BGMB	16.3	3.075	1.962	J
16	Cell Cycle	34	BGMB	7.3	1.039	0.868	J
17	Bioscience Reports	33	BGMB	6.4	0.73	0.856	J
18	Medical Science Monitor	33	M	5.2	0.601	0.86	J

Global Retracted Research Publications: A Bibliometric Study

19	Molecular Medicine Reports	33	M	5.9	0.65	0.721	J
20	Molecular Therapy Nucleic Acids	32	PTP	12.9	1.852	1.421	J
21	Applied Surface Science	31	PA	12.1	1.147	1.262	J
22	Molecular Biology Reports	28	BGMB	3.3	0.522	0.728	J
23	Biochemical And Biophysical Research Communications	27	BGMB	6.5	0.805	0.723	J
24	Frontiers In Oncology	27	M	4.5	1.291	1.191	J
25	Rsc Advances	27	CEGCE	5.9	0.667	0.833	J
26	Molecular Neurobiology	26	N	11	1.271	1.203	J
27	Life Sciences	25	PTP	8	1.132	1.199	J
28	Oncotargets And Therapy	25	M	7	0.811	0.812	J
29	Proceedings Of The National Academy Of Sciences Of The United States Of America	25	M	18.1	4.184	3.063	J
30	Nature	23	M	70.2	17.897	11.342	J
31	Construction And Building Materials	22	E	10.6	1.777	2.362	J
32	IEEE Access	22	E	6.7	0.927	1.326	J
33	Journal Of Nanoparticle Research	22	MCPA	3.6	0.417	0.527	J
34	Renewable And Sustainable Energy Reviews	21	E	28.5	3.678	4.535	J

M: Multidisciplinary; PA: Physics and Astronomy; BGMB: Biochemistry, Genetics and Molecular Biology; EPS: Earth and Planetary Sciences; PTP: Pharmacology, Toxicology and Pharmaceutics; CEGCE: Chemical Engineering: General Chemical Engineering; N: Neuroscience; E: Engineering; CS: Computer Science; MCPA: Math, Chemistry, Physics and Astronomy.

Table 7 : Source Type of Retracted Documents

Source Type	Number of Documents	%	Cummulative %
Conference Proceeding	11096	60.71	60.71
Journal	7159	39.17	99.89
Book Series	20	0.11	99.99
Undefined	1	0.01	100.00

was comparatively lesser than retracted documents in China. We have also found that both the authors with the highest number of retracted documents and retraction rate were from Islamic Azad University, Iran. Nazari, A has published 70 retracted documents. Ghoranneviss, M., had the highest rate of retraction, i.e., 24.65. It means that approximately every 25th document published by Ghoranneviss was retracted. The maximum number of retractions were observed in the conference proceedings, followed by Journal Articles. Both comprise 99.89% of the total retracted documents. The retracted documents were found in all type of sources. The retraction document appeared in high cite score as well as low cite score sources. However, the number of articles was comparatively higher in low cite score sources than in high cite score sources. Based on the analysis, it is recommended to have some strict worldwide guidelines for authors repeatedly publishing retraction works. We also recommend awareness and training programs for researchers to minimize the retraction of documents. The Librarians can take the initiative to organize such events in their respective institutions because “Librarians understand the nuances involved and can advocate for greater transparency around the retraction process and increase awareness of challenges posed by retractions”[10].

8. Future Work

The results of the present study showed that more than half of the retracted publications were from the conference proceedings. It will be interesting to explore, do these conference proceedings followed the peer review process or not? Other than these, the future work will be focused on citation patterns before or after the retraction, the time period between publication and retraction, and the significant reasons behind the retractions.

References

1. Feng, L., Yuan, J., & Yang, L. (2020). An observation framework for retracted publications in multiple dimensions. *Scientometrics*, 125(2), 1445-1457.
2. Samp, J. C., Schumock, G. T., & Pickard, A. S. (2012). Retracted publications in the drug literature. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, 32(7), 586-595.
3. Rosenkrantz, A. B. (2016). Retracted publications within radiology journals. *American Journal of Roentgenology*, 206(2), 231-235.
4. Rai, R., & Sabharwal, S. (2017). Retracted publications in orthopaedics: prevalence, characteristics, and trends. *JBJS*, 99(9), e44.
5. Wang, T., Xing, Q. R., Wang, H., & Chen, W. (2019). Retracted publications in the biomedical literature from open access journals. *Science and engineering ethics*, 25(3), 855-868.
6. Chauvin, A., De Villelongue, C., Pateron, D., & Yordanov, Y. (2019). A systematic review of retracted publications in emergency medicine. *European Journal of emergency medicine*, 26(1), 19-23.

7. Chambers, L. M., Michener, C. M., & Falcone, T. (2019). Plagiarism and data falsification are the most common reasons for retracted publications in obstetrics and gynaecology. *BJOG: An International Journal of Obstetrics & Gynaecology*, 126(9), 1134-1140.
8. Samuel, S., Cherian, J. M., & Thomas, A. M. (2020). Comprehensive analysis of retracted publications in dentistry: A 23-year review. *International Journal of dentistry*, 2020.
9. Zhang, Q., & Fu, H. Z. (2022). Productivity patterns, collaboration and scientific careers of authors with retracted publications in clinical medicine. *Scientometrics*, 127(4), 1883-1901.
10. Brown, S. J., Bakker, C. J., & Theis-Mahon, N. R. (2022). Retracted publications in pharmacy systematic reviews. *Journal of the Medical Library Association: JMLA*, 110(1), 47.
11. Kulkarni, A. V., Aziz, B., Shams, I., & Busse, J. W. (2009). Comparisons of citations in Web of Science, Scopus, and Google Scholar for articles published in general medical journals. *Jama*, 302(10), 1092-1096.
12. Rakhel, T. M., & Putera, P. B. (2021). Corruption in public procurement: A bibliometric analysis. *COLLNET Journal of Scientometrics and Information Management*, 15(2), 397-412.
13. [https://www.elsevier.com/__/data/assets/pdf_file/0007/69451/Scopus_Content_Coverage_Guide_WEB.pdf] accessed on 25 June 2022

Keywords: Bibliometrics; Bibliometric Analysis; Scientific Misconduct; Research Ethics; Retraction; Scopus

About Authors

Mr. Mohit Garg

Assistant Librarian

Central Library, Indian Institute of Technology Delhi, Delhi

Email: gargmohit@library.iitd.ac.in

Dr. Nabi Hasan

Librarian & Head

Central Library, Indian Institute of Technology Delhi, Delhi

Email: hasan@library.iitd.ac.in