

Global Landscape of Retracted Publications in Data Science: A Bibliometric Analysis (1993-2025)

Madhu S^{1,*}, Ashwini Kumar R², Aboobakar Siddiq³, Chaman Sab M⁴

¹Department of Library and Information Centre, GM University, Davangere, Karnataka, INDIA.

²Department of Library and Information Science, Kuvempu University, Shankaraghatta, Shimoga, INDIA.

³Department of Library and Information Science, Sadhana Collge (Base), Bangalore, Karnataka, INDIA.

⁴Department of Library and Information Centre, A.R.G. College of Arts and Commerce, College in Davangere, Karnataka, INDIA.

ABSTRACT

Background/Aim: This bibliometric study aims to present a global analysis of 5,398 retracted publications in Data Science from 1993 to 2025 using data from Retraction Watch, analysed with Microsoft Excel. **Methodology:** The study examines temporal trends, retraction delays, document types, nature of retraction, subject associations, country-wise production, and publisher and journals contributions. **Results:** Retractions were rare before 2015 but increased sharply after 2019, peaking in 2022 with 2,705 cases (50.11%). Most publications were retracted within one to two years (2,306; 42.72%). Technology (2,787) and Computer Science (2,215) were the most frequent co-associated subjects. China (3,902), India (962), and Saudi Arabia (304) accounted for the highest number of retractions. Hindawi (3,043) and IOS Press (1,161) were the leading publishers, research articles dominated document types (94.79%), and full retractions accounted for 98.72% of cases. **Conclusion:** The findings indicate that retractions in Data Science are a recent and global phenomenon, reflecting increased research output and strengthened editorial oversight.

Keywords: Bibliometrics, Data Science, Research Integrity, Retracted Publication, Retraction Watch, Scientific Misconduct.

Correspondence:

Dr. Madhu S

GM Central Library, GM University,
Davangere-577006, Karnataka, INDIA.
Email: madhu.rey24@gmail.com

Received: 16-12-2025;

Revised: 03-01-2026;

Accepted: 26-02-2026.

INTRODUCTION

Retractions serve as a fundamental corrective mechanism within scholarly communication, ensuring that flawed, unreliable, or unethical research is formally withdrawn from the scientific record to preserve the integrity and credibility of academic knowledge (Martini, 2016; Cao, 2017). Although comparatively rare relative to overall publication volume, the incidence of retracted articles has increased globally over recent decades, reflecting both the rapid expansion of research output and heightened scrutiny of research practices across disciplines (Fanelli, 2018). Retractions may result from honest errors, data fabrication, plagiarism, duplicate publication, or manipulation of the peer review process, and they play a crucial role in signalling to the research community and wider society that certain findings should no longer inform further enquiry or policy (Bar-Ilan and Halevi, 2017; Amit Kumar and Siwach, 2024). While most prior studies have focused on biomedical and clinical literature, there

is growing recognition of the importance of examining retraction patterns in fast-evolving, computationally intensive fields such as Data Science, where interdisciplinary applications and high publication volumes can amplify the impact of erroneous findings. The present study uses a bibliometric approach to systematically analyse global retracted publications in Data Science by examining temporal trends, subject associations, country contributions, publisher and journal patterns, and document types, thereby offering evidence-based insights into research integrity, editorial oversight, and the international landscape of scientific quality control (Madhu *et al.*, 2025).

REVIEW OF LITERATURE

Several studies have examined the patterns, causes, and consequences of retracted publications across disciplines, highlighting the growing concern over research integrity. Retractions are often considered corrective mechanisms to maintain the credibility of the scientific literature, yet their frequency and underlying reasons reveal systemic challenges in scholarly publishing (Gundur and Sampath Kumar, 2025). Examined 2,853 retracted publications authored by Indian scholars between 2010 and 2024. Their analysis indicated that a substantial proportion of retractions occurred during the period 2021-2024, accounting for 57.55% of the total. Fake peer



ScienScript

DOI: 10.5530/irc.3.1.4

Copyright Information :

Copyright Author (s) 2026 Distributed under
Creative Commons CC-BY 4.0

Publishing Partner : ScienScript Digital. [www.scienscript.com.sg]

review emerged as the leading cause of retraction, affecting 1,007 articles. Furthermore, the study identified eight major publishers including Springer, Elsevier, and Taylor and Francis accounting for 73.78% of all retracted publications.

Basumatary *et al.*, (2025) identified and studied the 79 retracted Indian business research articles published between 2014 and 2023 using data from the Retraction Watch and Scopus databases. The findings revealed a sharp increase in retractions, primarily attributed to fake peer review (46.83%) and plagiarism (31.64%). Tamil Nadu accounted for the highest number of retractions ($n=44$), while Springer was responsible for 41 cases, and the *Journal of Ambient Intelligence and Humanized Computing* alone contributed 17 retracted articles. Notably, most retractions occurred within one year of publication, suggesting increased editorial vigilance in addressing research misconduct. Kocyigit *et al.*, (2023) examined the 91 retracted publications from Kazakhstan, comprising 49 research articles (53.9%) and 39 conference papers (42.9%). The primary reasons for retraction were fake or biased peer review (38 papers, 41.8%), plagiarism (25 papers, 27.5%), and duplication (14 papers, 15.4%). In terms of international collaboration, Russia (24 papers, 26.4%) and China (5 papers, 5.5%) emerged as the leading partner countries.

Sharma, (2024) analysed 3,244 retracted publications by Indian researchers, revealing that 60% originated from private institutions, with fake peer review as the leading cause. Data integrity issues were more common in public and medical institutions, while plagiarism predominated in conference proceedings and non-Scopus journals. Domestic collaborations accounted for 80% of retractions, mainly in Q1-Q2 journals, whereas international collaborations (20%) were concentrated in Q1 journals (Amit and Siwach, 2023). Conducted a quantitative analysis of 93 Scopus-indexed retracted publications by Indian authors due to data falsification or fabrication, with an average retraction time of 2.62 years. The studies, mainly in Biochemistry, Genetics, and Molecular Biology, averaged 5.11 authors and 42.1 citations, with about half published as open access. Notably, citation rates showed little decline after retraction, underscoring the continued spread of flawed research and the need for stronger monitoring and ethical safeguards (Fang *et al.*, 2012). Reviewed 2,047 retracted biomedical and life-science articles indexed in PubMed and found that only 21.3% were retracted due to errors, while 67.4% resulted from misconduct, including fraud or suspected fraud (43.4%), duplicate publication (14.2%), and plagiarism (9.8%). The study reported a tenfold increase in fraud-related retractions since 1975 and identified distinct temporal and geographic patterns, highlighting systemic challenges in research integrity and reporting practices.

OBJECTIVES OF THE STUDY

The present study aims to undertake a bibliometric analysis of globally retracted publications in Data Science from 1993 to 2025, using data extracted from the Retraction Watch Database (RWD). The specific objectives are:

- To examine the global temporal distribution of retracted publications in Data Science and identify worldwide trends in retraction frequency over the years.
- To analyse retraction delays on a global scale, determining the time elapsed between publication and retraction across different countries and disciplines.
- To investigate the subject-wise and interdisciplinary distribution of retracted publications worldwide, highlighting the major domains associated with retractions in Data Science.
- To assess the global patterns of retractions by country, publisher, and journal, identifying key contributors to retracted literature and trends across regions.
- To evaluate the types of documents and the nature of retraction internationally, including full retractions, corrections, and expressions of concern, to understand worldwide editorial and post-publication practices in maintaining research integrity.

METHODOLOGY

This study employs a bibliometric research method to analyse global retracted publications in Data Science from 1993 to 2025. Bibliometric analysis facilitates a quantitative assessment of publication patterns, trends, and characteristics, thereby providing insights into research integrity and scholarly practices. The dataset was exclusively extracted from Retraction Watch (<https://retractiondatabase.org/>), a comprehensive and authoritative source that systematically tracks retracted publications worldwide. A total of 5,398 retracted publications were collected and categorised by year of publication, retraction delay, co-associated subjects, country, document type, publisher, journal, and nature of retraction. Microsoft Excel was used for data cleaning, classification, tabulation, and descriptive statistical analysis, including the calculation of frequency counts and percentages. This approach enabled the identification of temporal, disciplinary, geographical, and publisher-related trends. Overall, the bibliometric method, supported by spreadsheet-based analytical tools, allowed for a structured and large-scale examination of retraction dynamics, offering evidence-based insights into global patterns of retractions in Data Science.

Ethical Considerations

The study is based solely on publicly available data from Retraction Watch and does not involve human participants or

sensitive personal information. Ethical standards were rigorously maintained by ensuring the accuracy and reliability of extracted data, transparent reporting of sources, and objective interpretation of findings. All retracted publications are presented in aggregate form, and no individual authors or institutions are singled out, thereby respecting confidentiality and fairness. The research adheres to principles of responsible conduct, transparency, and integrity, consistent with international ethical guidelines for bibliometric and scientometric studies.

DATA ANALYSIS AND INTERPRETATION

Annual Retraction of Publications

The Figure 1 displays the year-wise analysis of retracted publications in Data Science, covering a total of 5,398 records globally, indicates a pronounced concentration of retractions from the period of 1993-2025. From 1993 to 2005, retractions were sporadic, with only one retracted publication each in 1993, 2001, 2003, and 2005, each contributing 0.02% to the total. A modest increase is observed between 2007 and 2008, with six retractions in 2007 (0.11%) and five in 2008 (0.09%), followed by a notable spike in 2009 with 36 retractions (0.67%). The subsequent years show fluctuations at a relatively low level, including six retractions in 2010 (0.11%), sixteen in 2011 (0.30%), nine in 2012 (0.17%), eight in 2013 (0.15%), and twelve in 2014 (0.22%). A clearer upward trend emerges from 2015 onwards, with twenty-four retractions in 2015 (0.44%), fifteen in 2016 (0.28%), forty-nine in 2017 (0.91%), and eighty-seven in 2018 (1.61%). The escalation becomes substantial in 2019 and 2020, recording 222 (4.11%) and 254 (4.71%) retractions respectively, before rising sharply in 2021 with 887 retractions (16.43%). The peak is observed in 2022, which alone accounts for 2,705 retractions, representing 50.11% of the total. Although a decline is evident

thereafter, retraction levels remain comparatively high in 2023 with 666 retractions (12.34%) and in 2024 with 384 retractions (7.11%), while the minimal count of three retractions in 2025 (0.06%) is likely attributable to incomplete data coverage. Overall, the statistical distribution clearly demonstrates that retractions in Data Science are predominantly a post-2019 phenomenon, reflecting intensified publication activity alongside strengthened post-publication review and research integrity enforcement mechanisms at the global level.

Frequency of Retraction Delays in Publications

The Figure 2 highlights the distribution of retraction delays for the 5,398 Data Science publications analysed indicates that the majority of retractions occur within a relatively short period after publication. Nearly one-third of the articles, amounting to 1,593 publications (29.51%), were retracted within the same year, while the largest proportion, 2,306 publications (42.72%), experienced a delay of one to two years before retraction. Collectively, these categories account for over seventy percent of all retractions, suggesting that issues leading to retraction are generally identified at an early stage. Retractions occurring within two to three years comprise 642 publications (11.89%), further supporting the prevalence of short retraction timelines. Beyond this period, the frequency declines considerably, with delays of three to four years accounting for 179 publications (3.32%), four to five years for 193 publications (3.58%), and five to six years for 185 publications (3.43%). Retractions with longer delays are comparatively uncommon, with fewer than two% of cases occurring after six years and only marginal proportions extending beyond ten years. The small number of retractions recorded after fifteen years, representing 19 publications (0.35%), indicates that very long delays are exceptional. Overall, the pattern suggests that retraction practices in Data Science are largely timely, reflecting

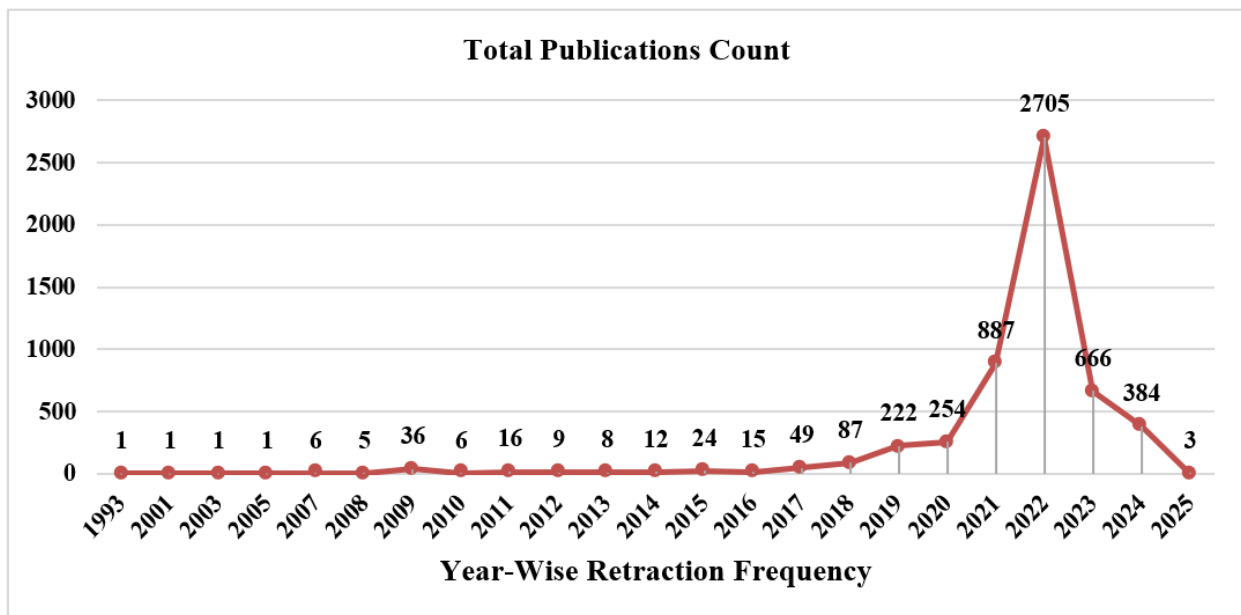


Figure 1: Year-Wise Research Outcome of Retracted Publications.

increased editorial oversight and more effective post-publication review mechanisms in contemporary scholarly publishing.

Document-Wise Retraction of Publications

Table 1 presents the distribution of document types among the 5,398 retracted publications in Data Science. Research articles overwhelmingly dominate, with 5,117 publications (94.79%), highlighting that primary research outputs are the most susceptible to post-publication correction. Conference abstracts and papers follow with 166 retractions (3.08%), reflecting the preliminary nature of conference-based dissemination. Review articles account for 43 retractions (0.80%), while clinical studies contribute 29 retractions (0.54%), indicating that secondary and applied research outputs are less frequently retracted than original research articles. The remaining document types, including letters, hybrid articles, preprints, meta-analyses, and editorials, each constitute less than 0.2% of the total, underscoring their minimal contribution to retracted literature. Overall, the distribution demonstrates that retractions in Data Science are predominantly associated with full-length research articles, emphasizing the critical importance of rigorous peer review and data validation at the primary research level.

Intra-Subject Association of Date Science Publications

Table 2 illustrates the distribution of co-associated subjects with retracted Data Science publications, comprising a total of 12,623 subject occurrences. Technology (2,787) and Computer Science (2,215) dominate the distribution, reflecting the core technological foundation of data science research. Among the

social sciences, Education records the highest association (779), followed by Sports and Recreation (376) and Psychology (198), indicating the increasing adoption of data-driven approaches in behavioral and educational studies. Business-related disciplines, including Business-Economics (374), Business-General (262), and Business-Management (244), also show substantial representation, highlighting the extensive use of data analytics in economic and managerial research. Health sciences contribute notably through Radiology and Imaging (296), Medicine-Oncology (202), Medicine-General (166), and Public Health and Safety (177), underscoring the integration of data science in medical and public health domains. Engineering, Environmental Sciences (165), Transportation (149), Linguistics (258), and Arts-General (143) further demonstrate the interdisciplinary reach of data science. The sizeable share of “Other 99 Subjects” (3,302) emphasizes the wide dispersion of retracted data science research across diverse academic fields, confirming the inherently interdisciplinary nature of the domain.

Country-Wise Distribution of Publications

Table 3 presents the country-wise distribution of retracted publications in Data Science, covering a total of 7,023 records. China accounts for the highest number of retractions with 3,902 publications, contributing a substantial share of the total, followed by India with 962 retractions. A second tier of countries includes Saudi Arabia (304), Pakistan (149), South Korea (142), and the United States (139), each reflecting notable levels of retraction activity. Several developing and emerging research economies, such as Ethiopia (123), Iraq (88), Malaysia (96), and Turkey (85), also feature prominently in the distribution. Among European

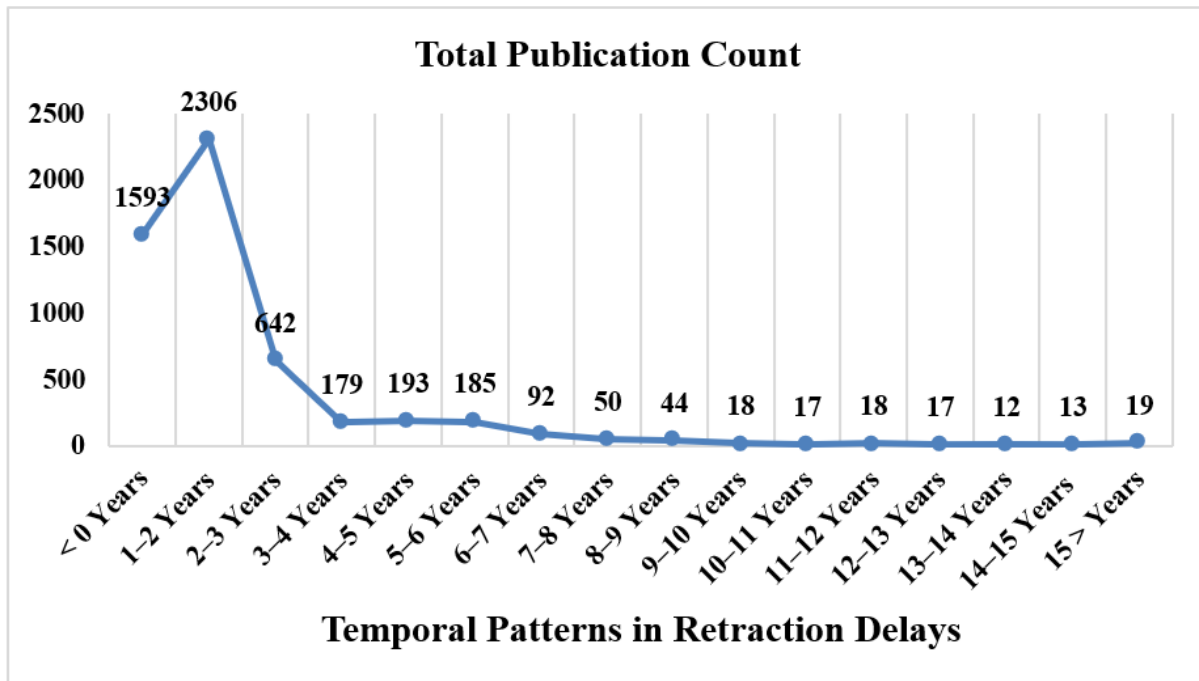


Figure 2: Retraction Delays in Publications.

countries, the United Kingdom records 97 retractions, while Spain accounts for 34. Other countries, including Egypt and Iran (53 each), Bangladesh (52), Taiwan (45), Jordan (35), Australia (34), and the Philippines (28), contribute smaller shares. The category comprising “Other 89 Countries”, with 602 retractions, indicates a wide global dispersion of retracted data science research. Overall, the distribution suggests that higher retraction counts are largely associated with countries having substantial research output in data science, rather than being confined to specific geographical regions, underscoring the global nature of research integrity challenges in this field.

Classification of Retracted Publications

Table 4 illustrates the nature of retracted publications in Data Science, encompassing a total of 5,398 records. The overwhelming majority, 5,329 publications (98.72%), are classified as retractions, indicating that complete withdrawal of the work from the scholarly record is the most common corrective measure. Expressions of concern account for 54 publications (1.00%), reflecting instances where potential issues are flagged without full retraction, often pending further investigation. Corrections are comparatively rare, comprising 15 publications (0.28%), which suggests that

Table 1: Document Type in retracted publications.

Sl. No.	Document Type	No. of Count	Percentage
1	Research Article	5117	94.79
2	Conference Abstract/Paper	166	3.08
3	Review Article	43	0.80
4	Clinical Study	29	0.54
5	Letter/Review Article	9	0.17
6	Article in Press/Research Article	8	0.15
7	Research Article/Review Article	5	0.09
8	Clinical Study/Research Article	3	0.06
9	Commentary/Editorial/Letter	3	0.06
10	Conference Abstract or Paper/Review Article	3	0.06
11	Preprint/Research Article	3	0.06
12	Meta-Analysis	2	0.04
13	Article in Press	1	0.02
14	Case Report	1	0.02
15	Conference Abstract or Paper/Meta-Analysis	1	0.02
16	Letter/Research Article	1	0.02
17	Meta-Analysis/Research Article	1	0.02
18	Meta-Analysis/Review Article	1	0.02
19	Retracted Article	1	0.02
Grand Total		5398	100.00

Table 2: Co-associated subjects with date science retracted publications.

Sl. No.	Co-associated Subjects	No. of Count	Sl. No.	Co-associated Subjects	No. of Count
1	(B/T) Technology	2787	11	(PHY) Engineering - Electrical	198
2	(B/T) Computer Science	2215	12	(SOC) Psychology	198
3	(SOC) Education	779	13	(SOC) Communications	181
4	(SOC) Sports and Recreation	376	14	(HSC) Public Health and Safety	177
5	(B/T) Business - Economics	374	15	(HSC) Medicine - General	166
6	(HSC) Radiology/Imaging	296	16	(ENV) Environmental Sciences	165
7	(B/T) Business - General	262	17	(PHY) Engineering - General	151
8	(SOC) Linguistics	258	18	(B/T) Transportation	149
9	(B/T) Business - Management	244	19	(HUM) Arts - General	143
10	(HSC) Medicine - Oncology	202	20	Other 99 Subjects	3302

minor errors or amendments constitute only a small fraction of post-publication interventions. Overall, the distribution demonstrates that most post-publication actions in Data Science involve full retractions, highlighting the field's emphasis on research integrity and the removal of flawed or unreliable work from the scientific record.

Journal-Wise Frequency of Retracted Publications

Table 5 presents the distribution of retracted publications in Data Science across journals, with a total of 5,398 records. The Journal of Intelligent and Fuzzy Systems lead with 1,159 retractions, representing the highest concentration of post-publication corrections in the field. It is followed by *Computational Intelligence and Neuroscience* with 497 retractions and *Security and Communication Networks* with 391, indicating that journals

Table 3: Top countries in date science retracted publications.

Sl. No.	Country	No. of Count	Sl. No.	Country	No. of Count
1	China	3902	11	Turkey	85
2	India	962	12	Egypt	53
3	Saudi Arabia	304	13	Iran	53
4	Pakistan	149	14	Bangladesh	52
5	South Korea	142	15	Taiwan	45
6	United States	139	16	Jordan	35
7	Ethiopia	123	17	Australia	34
8	United Kingdom	97	18	Spain	34
9	Malaysia	96	19	Philippines	28
10	Iraq	88	20	Other 89 Countries	602

Table 4: Nature of Retracted Publications.

Sl. No.	Retraction Type	No. of Count	Percentage
1	Retraction	5329	98.72
2	Expression of concern	54	1.00
3	Correction	15	0.28
Grand Total		5398	100.00

Table 5: Journal-Wise retraction publications count.

Sl. No.	Journal Name	No. of Count	Sl. No.	Journal Name	No. of Count
1	Journal of Intelligent and Fuzzy Systems	1159	11	Mathematical Problems in Engineering	136
2	Computational Intelligence and Neuroscience	497	12	Complexity	95
3	Security and Communication Networks	391	13	BioMed Research International	86
4	Wireless Communications and Mobile Computing	368	14	Advances in Multimedia	76
5	Soft Computing	264	15	Applied Bionics and Biomechanics	69
6	Journal of Healthcare Engineering	256	16	International Transactions on Electrical Energy Systems	58
7	Mobile Information Systems	164	17	PLoS One	58
8	Journal of Sensors	157	18	Annals of Operations Research	57
9	Computational and Mathematical Methods in Medicine	145	19	The International Journal of Electrical Engineering and Education	54
10	Journal of Environmental and Public Health	143	20	Other 300 Journals	1165

Table 6: Publisher-wise Distribution of Retracted Publications.

Sl. No.	Publisher Name	No. of Count	Sl. No.	Publisher Name	No. of Count
1	Hindawi	3043	11	IOP Publishing	27
2	IOS Press (bought by Sage November 2023)	1161	12	Elsevier - Cell Press	25
3	Springer - Nature Publishing Group	412	13	MDPI	22
4	Springer	179	14	EDP Sciences	17
5	IEEE: Institute of Electrical and Electronics Engineers	108	15	AIP Publishing	16
6	SAGE Publications	73	16	Frontiers	16
7	Wiley	65	17	Oxford University Press	10
8	Elsevier	61	18	ASTM International	7
9	PLoS	58	19	Sciendo Publishing	5
10	Taylor and Francis	52	20	Other 25 Journals	41

with high output in computational and technology-oriented research contribute substantially to retracted literature. Other notable journals include *Wireless Communications and Mobile Computing* (368), *Soft Computing* (264), and *Journal of Healthcare Engineering* (256), reflecting the interdisciplinary application of data science in engineering and health domains. Journals such as *Mobile Information Systems* (164), *Journal of Sensors* (157), *Computational and Mathematical Methods in Medicine* (145), and *Journal of Environmental and Public Health* (143) further demonstrate the breadth of data science research across applied fields. The category “Other 300 Journals” contributes 1,165 retractions, highlighting a wide dispersion of retracted publications across additional outlets. Overall, the data indicate that while retractions are concentrated in a few high-output journals, they also occur across a broad range of specialized and interdisciplinary journals, reflecting the extensive reach and impact of data science research.

Publisher-Wise Retraction of Publications

Table 6 presents the distribution of retracted publications in Data Science across publishers, based on a total of 5,398 records. Hindawi leads by a substantial margin with 3,043 retractions, indicating that a significant proportion of post-publication corrections in data science are concentrated within this publisher’s portfolio. IOS Press (now part of Sage as of November 2023) follows with 1,161 retractions, while Springer-Nature Publishing Group accounts for 412 retractions, and Springer alone contributes 179. Other notable publishers include IEEE (108), SAGE Publications (73), Wiley (65), Elsevier (61), and PLoS (58), reflecting moderate levels of retraction activity. Smaller contributions are observed from Taylor and Francis (52), IOP Publishing (27), Elsevier-Cell Press (25), MDPI (22), and a range of other publishers, each accounting for fewer than 20 retractions. The category “Other 25 Journals” contributes 41

retractions, illustrating the presence of a wider but comparatively minor distribution among additional outlets. Overall, the data indicate that retractions in Data Science are heavily concentrated among a few publishers, highlighting the critical role of editorial practices, peer review standards, and post-publication oversight in maintaining research integrity.

DISCUSSION

The bibliometric analysis of 5,398 retracted publications in Data Science reveals a strong concentration of retractions in recent years, underscoring the evolving nature of research integrity in this rapidly expanding field. Retractions were minimal prior to 2015 but increased sharply after 2019, culminating in a peak in 2022 with 2,705 retractions, accounting for 50.11% of the total. This trend indicates that retractions in Data Science are predominantly a post-2019 phenomenon, likely driven by intensified publication output, increased interdisciplinary applications, and strengthened post-publication scrutiny.

Retraction delay analysis shows that corrective actions are generally timely, with 1,593 publications (29.51%) retracted within the same year and 2,306 (42.72%) within one to two years, together comprising over 72% of all cases. Delays exceeding five years are uncommon, representing less than 7% of retractions, reflecting improved editorial vigilance and research integrity enforcement.

Research articles constitute the overwhelming majority of retractions (5,117; 94.79%), highlighting the vulnerability of primary research to methodological and ethical lapses. Subject-wise analysis confirms the interdisciplinary reach of Data Science, with Technology (2,787) and Computer Science (2,215) dominating, followed by Education (779), Business-related disciplines (880 combined), and Health Sciences, including Radiology/Imaging (296) and Medicine-Oncology (202).

Geographically, retractions are concentrated in high-output research countries, led by China (3,902) and India (962), together accounting for over two-thirds of country-linked retractions. At the publishing level, Hindawi (3,043) and IOS Press (1,161) collectively contribute nearly 78% of all retractions, while journal-level analysis shows notable clustering within a small number of high-volume outlets. Finally, the predominance of full retractions (5,329; 98.72%) over corrections or expressions of concern reflects a strong preference for complete withdrawal when significant flaws are identified. Collectively, these patterns highlight retractions as an increasingly central mechanism for safeguarding the integrity of global Data Science research.

CONCLUSION

The comprehensive analysis of 5,398 retracted publications in Data Science demonstrates that retractions are predominantly a recent and concentrated phenomenon, with the majority occurring post-2019 and peaking in 2022, reflecting the rapid expansion of research output in the field. Most retractions are identified within one to two years of publication, underscoring the effectiveness of contemporary post-publication monitoring and editorial oversight. The data reveal that retractions are highly interdisciplinary, spanning Technology, Computer Science, Education, Business, Health Sciences, and Engineering, highlighting the pervasive application of data-driven research across diverse domains. High-output countries, particularly China, India, and Saudi Arabia, and a limited set of publishers and journals contribute disproportionately to retracted literature, indicating a correlation between publication volume and retraction frequency. Research articles constitute the overwhelming majority of retractions (94.79%), and full retractions (98.72%) dominate corrective actions, emphasizing the field's commitment to maintaining research integrity. Overall, these findings underscore the critical importance of rigorous peer review, ethical research practices, and vigilant editorial mechanisms in safeguarding the reliability and credibility of Data Science scholarship worldwide.

ABBREVIATIONS

RWD: Retraction Watch Database; **IEEE:** Institute of Electrical and Electronics Engineers; **MDPI:** Molecular Diversity Preservation International; **IOP Publishing:** Institute of Physics Publishing; **EDP Sciences:** Édition Diffusion Presse Sciences; **ASTM International:** American Society for Testing and Materials

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCES

- Bar-Ilan, J., and Halevi, G. (2017). Post retraction citations in context: a case study. *Scientometrics*, 113(1), 547-565. <https://doi.org/10.1007/s11192-017-2242-0>
- Basumatary, B., Sinha, M. K., Kumar, V., and Verma, M. K. (2025). Retractions in Indian Business Research: A Closer Look at Trends, Reasons and Impacts of Retracted Articles. *DESIDOC Journal of Library and Information Technology*, 45(2), 131-140. <https://doi.org/10.14429/djlit.45.2.20357>
- Cao, L. (2017). Data science: Challenges and directions. *Communications of the ACM*, 60(8), 59-68. <https://doi.org/10.1145/3015456>
- Fanelli, D. (2018). Is science really facing a reproducibility crisis, and do we need it to? *Proceedings of the National Academy of Sciences*, 115(11), 2628-2631. <https://doi.org/10.1073/pnas.1708272114>
- Fang, F. C., Steen, R. G., and Casadevall, A. (2012). Misconduct accounts for the majority of retracted scientific publications. *Proceedings of the National Academy of Sciences*, 109(42), 17028-17033. <https://doi.org/10.1073/pnas.1212247109>
- Gundur, B. N., and Sampath Kumar, B. T. (2025). Trajectory of Research Retractions by Indian Scholars: Trends, Causes and Ethical Imperatives. *Journal of Data Science, Informetrics, and Citation Studies*, 4(2), 257-262. <https://doi.org/10.5530/jcitation.20250207>
- Kocygigit, B. F., Zhaksylyk, A., Akyol, A., and Yessirkepov, M. (2023). Characteristics of Retracted Publications from Kazakhstan: An Analysis Using the Retraction Watch Database. *Journal of Korean Medical Science*, 38(46), 1-11. <https://doi.org/10.3346/jkms.2023.38.e390>
- Kumar, A., and Siwach, A. K. (2023). Analysis of Retracted Publications by Indian Authors due to Falsification or Fabrication of Data. *Kelpro Bulletin*, 27(2), 29-44.
- Kumar, A., and Siwach, A. K. (2024). Analysis of Indian Retracted Publications: A Study Based on Scopus Data. *The Serials Librarian*, 85(7-8), 191-204. <https://doi.org/10.1080/00361526X.2024.2306396>
- Madhu, S., Kannappanavar, B. U., and Sab M, C. (2025). Scientometric Portrait of India's Renowned Chemist from Indian Institute of Science: Dr. Gautam Radhakrishna Desiraju. *Journal of Data Science, Informetrics, and Citation Studies*, 4(1), 47-53. <https://doi.org/10.5530/jcitation.20250170>
- Martini, B. (2016). The Data Revolution. *Big Data, Open Data, Data Infrastructures and Their Consequences*. *Regional Studies*, 50(3). <https://doi.org/10.1080/00343404.2015.1107987>
- Sharma, K. (2024). Over two decades of scientific misconduct in India: Retraction reasons and journal quality among inter-country and intra-country institutional collaboration. *Scientometrics*, 129(12), 7735-7757. <https://doi.org/10.1007/s11192-024-05192-z>

Cite this article: Madhu S, Kumar AR, Siddique A, Sab CM. Global Landscape of Retracted Publications in Data Science: A Bibliometric Analysis (1993-2025). *Info Res Com.* 2026;3(1):34-41.