J Data Sci. Info. Citation Studies., 2024; 3(1):1-10.
https://www.jcitation.org

Research Article

Unveiling the Status of Open Access Dairy Research in India through Data Carpentry

Abhijit Roy, Akhandanand Shukla*

[1]Department of Library & Information Science, Central University of Tamil Nadu, Thiruvarur, Tamil Nadu, INDIA.
[2]Department of Library & Information Science, North-Eastern Hill University, Shillong, Meghalaya, INDIA.

ABSTRACT

The present study examined the open-access status of dairy research publications published by Indian institutions. The study utilized Scopus to source the required data for a 10-year period (2013-2022) and employed a data carpentry approach for cleaning and processing the data. Data carpentry tools/software such as Open Refine, OpenAlex, Dimensions.ai, and Altmetric.com were utilized in various instances for different purposes. Findings reveal that 35.26% of publications were distributed across various open-access categories and were used in the study to generate the analysis. The study indicates that DOI-based publications increased over time, while Open-Access (OA) publications consistently maintained a share exceeding 30%. Gold OA is prevalent in terms of occurrence, citations, and altmetrics, with other OA categories also significantly present. The CC-BY license is mostly used as an OA license, and among the 10 document types, "Article" is the most prevalent. The "Journal of Food Science and Technology" holds the highest number of OA publications, while Elsevier BV is the highest contributing publication agency. Further study reveals that the three-authorship pattern mostly occurred in open-access publications of dairy research, with a prevalent multiple-authorship pattern. The ratio of citations was found to be higher for open-access publications compared to closed access, and the altmetric attention score was also found to be higher for open access publications than for closed ones. ICAR-National Dairy Research Institute emerged as the top contributor of OA publications in dairy research. Lastly, the study discusses the limitations and future prospects of the research, providing a comprehensive overview of the conducted analysis.

Keywords: Open Access Publication, Dairy Research, OpenAlex, OpenRefine, Data Carpentry.

Correspondence:
Akhandanand Shukla
Department of Library & Information Science, North-Eastern Hill University, Shillong - 793002, Meghalaya, INDIA.
Email: akhandanandshukla@gmail.com

Received: 19-01-2024;
Revised: 19-01-2024;
Accepted: 19-01-2024.

INTRODUCTION

The concept of open access has led to an evolution in scholarly publications. According to Suber (2012), "Open Access (OA) literature is digital, online, free of charge, and free of most copyright and licensing restrictions." In the realm of scholarly publications, two popular OA models exist for scholars: the Green path and the Gold path. Green OA involves publications in closed access journals but preserves them in archives or repositories, while Gold OA entails publishing in open access journals with an attached open license (Piwowar *et al*, 2018). In 2002, the scenario of open access publications underwent a significant change with the adoption of the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities on an international level (Chan *et al.*, 2002).

India has also embraced open access policies and enjoys the benefits of citations and altmetric data (Fernandez, 2006). Open access publications, however, face various challenges and barriers as outlined by Björk (2004). Singh *et al.* (2020) assessed the Open Access (OA) models (Gold and Green) in their study from an Indian perspective and found that 24% of Indian publications are adopting open access routes. Similar studies have been conducted to assess open access levels globally (Archambault *et al.*, 2013; Archambault *et al.*, 2014) in specific regions such as Europe (Maddi *et al.*, 2021) using the search engine Google (Martín-Martín *et al.*, 2018) across various subject disciplines (Piwowar *et al.*, 2018; Piwowar *et al.*, 2019) and at universities worldwide (Robinson-Garcia *et al.*, 2019; Robinson-Garcia *et al.*, 2020).

Furthermore, Bhat (2009) conducted a study to assess the open access status of premier Indian institutions, while Chakravarty and Mahajan (2011) and Keisham and Sophiarani (2008) revealed the development of open access journals in India in their respective studies. In his research, Nazim (2021) explored the status of open access publications in Indian institutions. Additionally, Mukhopadhyay (2022) devised an open access friendliness

DOI: 10.5530/jcitation.3.1.1

ranking framework system, which measured the open access friendliness of IITs through data carpentry. Various researchers studied the open access friendliness of Indian higher education institutions (Roy and Mukhopadhyay, 2022), central universities (Roy and Mukhopadhyay, 2022), state universities (Roy and Mukhopadhyay, 2022), and NITs (Roy and Mukhopadhyay, 2022).

Despite being a developing country with an economy largely based on agriculture, India has initiated numerous endeavors supported by the Government of India (Agrawal, 1980). Various agriculture institutions, information centers, and agriculture libraries have been established by the central and state governments for research and education purposes. Researchers such as Arunachalam (1998), Arunachalam and Umarani (2001), and Sagar *et al.* (2014) have explored agriculture research in India, while Raychoudhury and Sarkhel (2011) have examined agriculture research in West Bengal using bibliometric and scientometric approaches. Additionally, several researchers (Lal *et al.*, 1996; Sangam *et al.*, 2014; Suryanarayana, 2002; Tripathi *et al.*, 2013) have attempted to assess the publication status of various disciplines within the agriculture domain. Several authors (Abraham *et al.*, 2009; Guttikonda *et al.*, 2009; Kumar *et al.*, 2012; Shukla and Singh, 2009) have discussed open access agriculture research publications in Indian open access journals.

Given this backdrop, the present study delves into the open access status of dairy research publications from an Indian perspective. The study employs the open-source tool OpenRefine to gather open access status from OpenAlex, citation data from Dimension.ai, and altmetric data from Altmetric.ai. Data carpentry approaches have been utilized to achieve the objectives.

OBJECTIVES

The main aim of the study is to reveal the status of open access publications in Indian Dairy Research. To achieve this, the study has the following specific objectives:

To analyze the growth of open access publications over the past decade (2013-2022).

To investigate the utilization of licenses in open access publications over the past decade.

To categorize the types of documents within the field of dairy research.

To identify the leading journals and publishers in the realm of open access publications.

To examine the authorship patterns in open access publications.

To assess the citation patterns for open access publications.

To analyze the altmetric attention scores for dairy research.

To identify the top-performing Indian institutions in terms of open access.

METHODOLOGY

Data Collection

The current study has investigated the status of open access publications in Indian dairy research over the past decade. For this purpose, the Scopus database was chosen as the primary data source. In this study, specific dairy research related keywords, such as dairy, dairy farming, dairy product, dairy foods, dairy cattle, dairy technology, dairy science, dairy processing and dairy farm were selected and filtered by affiliation country (India). The study encompassed the extraction of data from Scopus for the last ten years (2013-2022). The following search string was used for data collection:

TITLE-ABS-KEY ("Dairy") OR TITLE-ABS-KEY ("Dairy farming") OR TITLE-ABS-KEY ("dairy product") OR TITLE-ABS-KEY ("Dairy foods") OR TITLE-ABS-KEY ("Dairy Cattle") OR TITLE-ABS-KEY ("dairy technology") OR TITLE-ABS-KEY ("dairy science") OR TITLE-ABS-KEY ("Dairy science and technology") OR TITLE-ABS-KEY ("dairy chemistry") OR TITLE-ABS-KEY ("Dairy processing") OR TITLE-ABS-KEY ("dairy farm") AND (LIMIT-TO (AFFILCOUNTRY, "India")) AND (LIMIT-TO (PUBYEAR, 2013) OR LIMIT-TO (PUBYEAR, 2014) OR LIMIT-TO (PUBYEAR, 2015) OR LIMIT-TO (PUBYEAR, 2016) OR LIMIT-TO (PUBYEAR, 2017) OR LIMIT-TO (PUBYEAR, 2018) OR LIMIT-TO (PUBYEAR, 2019) OR LIMIT-TO (PUBYEAR, 2020) OR LIMIT-TO (PUBYEAR, 2021) OR LIMIT-TO (PUBYEAR, 2022).

Data Mining

The present study collected a total of 4,597 documents from the Scopus database. Subsequently, we conducted a comprehensive assessment of the publication datasets using open-source software OpenRefine. Figure 1 visually represents the annual publication status, distinguishing between those with DOI and those without DOI, in the field of dairy research in India. OpenRefine yielded 3,571 publications that possessed a DOI. These 3,571 DOI-associated publications underwent further scrutiny by sending queries to three different databases to ascertain their open access status (OpenAlex), citation counts (Dimensions.ai), and altmetric attention scores (altmetric.com). The responses received from these databases are given in Table 1. It's important to note that out of the total publications (4,597), 1,026 were without DOI and were consequently excluded from the analysis, as DOI status is closely linked to open access status. Additionally, among the publications with DOIs, 32 were found to have unknown open access status in OpenAlex and were also excluded from the open access status analysis.

Open Access Status: This study obtained information on the open access status of publications from OpenAlex (Priem *et al.*, 2022).

Table 1: **Use of REST API for data fetching from different databases.**

Queries Sent	REST API	Responses received
3,571 Publications with DOI	https://api.openalex.org/sources?search= + DOI	3,539 (99.10%)
	https://metrics-api.dimensions.ai/doi/ + DOI	3,502 (98.07%)
	https://api.altmetric.com/v1/doi/ + DOI	882 (24.70%)

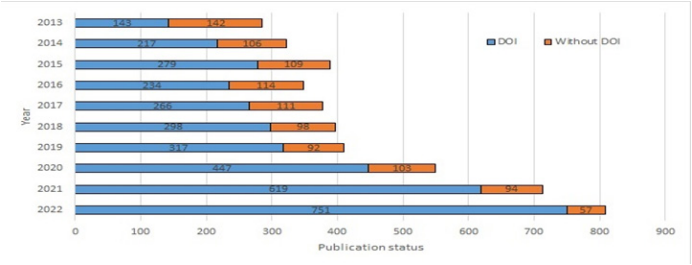

Figure 1: Publication's status of dairy research in India (2013-2022).

We received 3,539 responses regarding the open access status out of 3,571 DOI publications. This was achieved through a REST API call using OpenRefine, as outlined in Table 1.

Citation Status: Data related to the citation of open access publications was retrieved from Dimensions.ai. A total of 3,571 DOI requests were submitted to obtain the citation status, and we received 98.07% (*n*=3,502) responses from Dimensions.ai.

Altmetric Attention Score: REST API calls were made for a total of 3,571 publications with DOIs to altmetric.com to obtain the altmetric attention score. Out of the total requests sent, we received responses for 882 publications, accounting for 24.7% of the total requests.

DATA ANALYSIS AND DISCUSSION

Growth of Open Access Publications

The Scopus database yielded 4,597 documents on dairy research, which were processed with OpenRefine to determine the DOI and non-DOI status of publications. OpenRefine identified 3,571 publications with DOI and 1,026 publications without DOI. As a result, a significant portion of publications (22.32%, *n*=1,026 out of 4,597 publications) was excluded from the study. Figure 1 provides an overview of the publication status with DOI over the span of ten years. It highlights that in 2013, the highest number of research papers, totaling 142 (49.82%), were published without DOI. However, as time progressed, the proportion of publications with DOI increased to 92.94%, while those without DOI decreased to 7%.

Table 2 presents the status of open and closed access publications obtained from OpenAlex. It also illustrates the annual growth of closed and open access publications. Among the 3,539 publications to which OpenAlex responded, 64.74% (*n*=2,291) were categorized as closed access, while 35.26% (*n*=1,248) were categorized as open access. The open access publications were further categorized into four distinct OA routes. Among these, the majority of publications (64.26%, *n*=802 out of the total open access publications) were distributed as Gold OA, followed by 17.23% (*n*=215 out of the total open access publications) in Green OA, 9.46% (*n*=118 out of the total open access publications) in Bronze OA, and 9.05% (*n*=113 out of the total open access publications) in Hybrid OA. Table 2 also indicates that over the past ten years, the ratio of open access publications consistently exceeded 30%. The highest proportion of open access publications (47.14%) was observed in 2013.

Licenses in Open Access Publications

In the licensing scenario, 57.77% (*n*=721 out of the total OA publications) of dairy research publications were published with a license, while 42.29% (*n*=527) of OA publications were published without any license. Among the publications with licenses, ten types of licenses were found attached to these OA publications. These licenses included CC-BY (481 publications, constituting 38.54% of the total OA documents), CC-BY-NC-ND (110 publications, 8.81%), CC-BY-NC-SA (50 publications, 4%), CC-BY-NC (45 publications, 3.61%), Elsevier-Specific: OA User License (23 publications, 1.84%), CC-BY-SA (6 publications, 0.48%), CC0 (3 publications, 0.24%), CC-BY-ND (1 publication, 0.08%), Implied-OA (1 publication, 0.08%), and Public Domain (1 publication, 0.08%).

Upon reviewing Table 3, it becomes evident that a significant portion of Green OA publications, specifically 88.37%, have been published without licensing. This is followed by Bronze OA (80.51%), Gold OA (29.93%), and Hybrid OA (1.77%). Interestingly, Hybrid OA has the highest proportion of licensed publications, accounting for 98.23% (*n*=111 publications), followed by Gold OA at 70.07%.

Document Types in Open Access

In scholarly communications, document types play a crucial role, and their examination is a focal point of this study. Table 4 presents a comprehensive overview of various document types and their respective proportions within dairy research publications. A total of ten distinct document types were identified. Upon reviewing Table 4, it is evident that the "Article" document type constitutes the majority, accounting for 82.69% of the total OA publications, followed by "Review" at 14.02%, "Conference Paper" at 1.28%, and "Book" at 0.56%. Other document types appeared with significantly lower frequencies. It's worth noting that across various open access categories, including Green OA (86.51%), Gold OA (80.67%), Hybrid OA (84.07%), and Bronze OA (88.14%), the majority of publications take the form of "Article," followed by "Review" types.

Table 2: Open access publications in dairy research (2013-2022).

Year	Total Pub.	Closed Access Pub.	Open Access Pub.	OA%	Open Access Routes			
					Green OA	Gold OA	Hybrid OA	Bronze OA
2013	140	74	66	47.14%	11	35	7	13
2014	216	128	88	40.74%	38	34	11	5
2015	279	190	89	31.90%	36	40	7	6
2016	233	152	81	34.76%	16	39	16	10
2017	262	173	89	33.97%	14	56	8	11
2018	296	196	100	33.78%	20	67	3	10
2019	316	203	113	35.76%	9	81	12	11
2020	445	304	141	31.69%	14	101	10	16
2021	616	400	216	35.06%	27	155	21	13
2022	736	471	265	36.01%	30	194	18	23
Total	3,539	2,291	1,248	---	215	802	113	118

Table 3: Use of license in OA publications in dairy research (2013-2022).

License	License Type	Green OA	Gold OA	Hybrid OA	Bronze OA	Total
With License	Public Domain (PD)	1 (0.47%)	--	--	--	1 (0.08%)
	CC-BY	14 (6.51%)	410 (51.12%)	57 (50.44%)	--	481 (38.54%)
	CC-BY-SA	--	5 (0.62%)	1 (0.88%)	--	6 (0.48%)
	CC0	1 (0.47%)	2 (0.25%)	--	--	3 (0.24%)
	CC-BY-NC	3 (1.40%)	25 (3.12%)	17 (15.04%)	--	45 (3.61%)
	CC-BY-NC-ND	4 (1.86%)	73 (9.10%)	33 (29.20%)	--	110 (8.81%)
	CC-BY-NC-SA	2 (0.93%)	46 (5.74%)	2 (1.77%)	--	50 (4%)
	CC-BY-ND	--	1 (0.12%)	--	--	1 (0.08%)
	Elsevier-Specific: OA User License	--	--	--	23 (19.49%)	23 (1.84%)
	IMPLIED-OA	--	--	1 (0.88%)	--	1 (0.08%)
Without License		190 (88.37%)	240 (29.93%)	2 (1.77%)	95 (80.51%)	527 (42.29%)
Total		215 (100%)	802 (100%)	113 (100%)	118 (100%)	1248 (100%)

Leading Journals and Publishers in Open Access

The study primarily focused on scholarly communications published by open access journals, and the relevant data is presented in Table 5. Table 5 showcases the top ten open access journals within the field of dairy research, collectively accounting for 33.97% of open access publications out of the total OA publications. It's important to note that a total of 427 journals contributed to the 1,248 open access publications. Among these, within Green OA publications, 59.53% of documents were

Table 4: Different types of documents published in diary research.

Document Types	Open Access Pub.	Open Access Categories			
		Green OA	Gold OA	Hybrid OA	Bronze OA
Article	1032 (82.69%)	186 (86.51%)	647 (80.67%)	95 (84.07%)	104 (88.14%)
Review	175 (14.02%)	18 (8.37%)	132 (16.46%)	15 (13.27%)	10 (8.47%)
Book Chapter	2 (0.16%)	1 (0.47%)	1 (0.12%)	--	--
Conference Paper	16 (1.28%)	--	13 (1.62%)	1 (0.88%)	2 (1.69%)
Book	7 (0.56%)	7 (3.26%)	--	--	--
Short Survey	3 (0.24%)	--	2 (0.25%)	1 (0.88%)	--
Editorial	6 (0.48%)	--	4 (0.50%)	1 (0.88%)	1 (0.85%)
Note	5 (0.40%)	2 (0.93%)	2 (0.25%)	--	1 (0.85%)
Letter	1 (0.08%)	1 (0.47%)	--	--	--
Data Paper	1 (0.08%)	--	1 (0.12%)	--	--
Total	1,248 (100%)	215 (100%)	802 (100%)	113 (100%)	118 (100%)

published by the top ten OA journals, while 31.55% of documents were published in Gold OA, 28.81% in Bronze OA, and 7.96% in Hybrid OA. Notably, the "*Journal of Food Science and Technology*" holds the highest number of publications (*n*=118) in dairy research, followed by "*Veterinary World*" (*n*=96) and "*Journal of Food Processing and Preservation*" (*n*=45). Of the top ten OA journals, six belong to India, with the remaining three originating from the USA (*n*=3) and China (*n*=2).

In total, 136 publication agencies have contributed to the publication of 1,248 open access documents. The study identifies the top ten publishers and their respective contributions, as outlined in Table 6. These top ten publication agencies collectively account for 51.76% (*n*=646) of the total OA publications. The highest contribution comes from Elsevier BV (*n*=120, constituting 9.62% of total OA publications), followed by the European Bioinformatics Institute (119, 9.54%), Veterinary World (95, 7.61%), etc. Among the total OA publications, the highest share was associated with Gold OA (*n*=802), followed by Green OA (*n*=215), Bronze OA (*n*=118), and Hybrid OA (113). Within these OA categories and among the top ten publication agencies, the highest share is attributed to Green OA (76.28% of total Green OA), followed by Gold OA (49.25% of total Gold OA), Hybrid OA (46.9% of total Hybrid OA), and Bronze OA (28.81% of total Bronze OA).

Table 6 also provides insights into the citation share of OA publications based on publication agencies, with Elsevier BV receiving the highest number of citations (3804), followed by the European Bioinformatics Institute (3115) and Springer Sciences Business Media (2119). The contributions of the top ten publication agencies collectively account for 27.09% of the total OA citations. In terms of Altmetric scores, the top-scoring publication agency was Frontiers Media (960.4), followed by Elsevier BV (904.22) and the European Bioinformatics Institute (628.25). Like the citation share of top ten publishers, the Altmetric score of these top ten publishers amounts to 4053.25 (27.08% of the total Altmetric Score of OA publications). Additionally, a total of 727 licenses were attached to OA publications, with 55.98% of these licenses associated with the top ten publication agencies. Among the top ten publication agencies, Elsevier BV has the highest number of licenses attached (111), followed by Veterinary World (95), MDPI (75), and Frontiers Media (47). Other publishers have relatively fewer licenses attached to their publications.

Authorship Patterns in Open Access Publications

The study examined authorship patterns in dairy research from 2013 to 2022. Over the past decade, 1,248 open access scholarly publications were contributed by 6,657 authors. Table 7 provides insights into the authorship patterns of open access research across different OA categories. Among OA publications, the highest number of papers (*n*=223, 17.87%) were authored by 3 authors, while the highest number of authors (*n*=1289) appeared in publications with more than 10 authors. Within the open access categories, 64.66% of papers (*n*=807) were authored by 1 to 5 authors, while 35.34% of papers (*n*=441) were authored by 6 to 10+ authors. Conversely, 41.58% of authors (*n*=2768) were accounted for 807 papers, while 58.41% of authors (*n*=3889) were accounted for 441 papers.

Among these OA categories, Gold OA had the highest number of papers (*n*=802, 64.26% of total OA publications) and contributors (*n*=4295, 64.51% of total authors), followed by Green OA (*n*=215, 1008 authors), Bronze OA (*n*=118, 634 authors), and Hybrid OA (*n*=113, 720 authors). In the case of Gold OA, 62.46% of papers (*n*=501) were authored by 1 to 5 authors, while 37.54% of papers (*n*=301) were authored by 6 to 10+ authors. Conversely, 60.46% of authors (*n*=2597) were accounted for 301 papers, while 39.54% of authors (*n*=1698) were accounted for 501 papers. A similar trend was found for Hybrid OA and Bronze OA, while Green OA exhibited a reverse trend.

Citation Patterns in Open Access Publications

Table 8 displays the citation patterns of dairy research over the past ten years (2013-2022), specifically highlighting the citation data for closed access publications for comparison with open access publications. The citation data, collected from Dimensions.ai in June 2023, includes responses from 98.07% (*n*=3,502) out of 3,571 publications with DOIs sent for citation status. Six documents did not receive a response regarding closed and open access status and were therefore excluded from the study. Consequently, a total of 3,496 publications received 51,071 citations, with 19.28% (*n*=674) not receiving any citations. Out of the total publications, 2,249 belonged to closed access, receiving 30,261 citations, accounting for 59.25% of total citations (Figure 2). In closed access, the citation per paper ratio was 13.46. Conversely, in the case of open access, a total of 20,810 citations (40.75%) were received by 1,247 OA publications, resulting in a citation ratio of 16.69. In the field of dairy research, the citation rate of open access publications was found to be better than closed access.

Within the realm of open access categories, the highest citation ratio was found for Hybrid OA (29.23), followed by Green OA (19.97), Gold OA (14.46), and Bronze OA (13.89). Among the OA categories, the highest share of citations (55.71%) was found for Gold OA, covering 64.31% of documents, followed by Green OA (20.54% citations share, 17.16% documents), Hybrid OA (15.87% citations share, 9.06% documents), and Bronze OA (7.87% citations share, 9.46% documents). The highest number of citations (582) originated from a hybrid OA publication affiliated with The Vallabhbhai Patel Chest Institute (VPCI), bearing a

Table 5: **Top ten open access journals and their contributions.**

Journals Name and Country	Total OA Pubs.	Green OA	Gold OA	Hybrid OA	Bronze OA
Journal of Food Science and Technology (IN)	118	116	0	0	2
Veterinary World (IN)	96	1	95	0	0
Journal of Food Processing and Preservation (USA)	45	0	45	0	0
Asian Journal of Dairy and Food Research (IN)	42	0	42	0	0
Indian Journal of Animal Sciences (IN)	25	0	24	1	0
Indian Journal of Animal Research (IN)	19	1	0	0	18
Journal of Dairy Science (USA)	17	0	1	2	14
PLoS ONE (USA)	17	0	17		
Journal of Pure and Applied Microbiology (IN)	16	0	14	2	0
Frontiers in Microbiology (CH)	15	0	15	0	0
3 Biotech (CH)	14	10	0	4	
Total	424 (33.97%)	128 (59.53%)	253 (31.55%)	9 (7.96%)	34 (28.81%)
Others Journals	824 (66.03%)	87 (40.47%)	549 (68.45%)	104 (92.04%)	84 (71.19%)

Table 6: Top ten open access publisher and their contributions.

Publisher Name	OA Articles	Green OA	Gold OA	Hybrid OA	Bronze OA	Citations	Altmetric Score	License
Elsevier BV	120	2	58	37	23	3804	904.22	111
European Bioinformatics Institute	119	118	1	0	0	3115	628.25	4
Veterinary World	95	0	95	0	0	1761	122.2	95
MDPI	78	0	78	0	0	775	500.96	75
Wiley-Blackwell	61	0	57	4	0	687	69.11	6
Frontiers Media	47	0	47	0	0	1116	960.04	47
National Institutes of Health	38	38	0	0	0	294	48.68	0
Springer Science Business Media	37	5	11	11	10	2119	504.85	22
Medknow	27	1	25	0	1	153	311.95	23
Indian Council of Agricultural Research	24	0	23	1	0	10	3	24
	646 (51.76%)	164 (76.28%)	395 (49.25%)	53 (46.90%)	34 (28.81%)	13,834 (27.09%)	4,053.25 (27.08%)	407 (55.98%)
Other Publishers	602	51	407	60	84	37,237	10,914.81	320
Total	1,248	215	802	113	118	51,071	14,968.06	727

Table 7: Authorship patterns in Dairy research.

No. of Author(s)	Open Access		OA Categories							
			Green OA		Gold OA		Hybrid OA		Bronze OA	
	NP	AU	NP	AU	NP	AU	NP	AU	NP	AU
1	25	25	5	5	13	13	2	2	5	5
2	169	338	31	62	113	226	10	20	15	30
3	223	669	48	144	147	441	19	57	9	27
4	214	856	45	180	122	488	21	84	26	104
5	176	880	34	170	106	530	15	75	21	105
6	140	840	14	84	96	576	13	78	17	102
7	92	644	12	84	60	420	11	77	9	63
8	59	472	5	40	43	344	6	48	5	40
9	36	324	6	54	23	207	3	27	4	36
10	32	320	5	50	24	240	3	30	0	0
10+	82	1289	10	135	55	810	10	222	7	122
Total	1,248	6,657	215	1,008	802	4,295	113	720	118	634

NP=Number of Papers, AU=Authors.

Table 8: **Pattern of citation.**

Citation Range	Closed Access		Open Access		OA Categories							
					Green OA		Gold OA		Hybrid OA		Bronze OA	
	NP	TC	NP	TC	NP	TC	NP	TC	NP	TC	NP	TC
0	456	0	218	0	12	0	166	0	16	0	24	0
1-10	1070	4360	589	2573	101	492	388	1702	46	162	54	217
11-20	303	4552	208	3148	52	788	120	1806	22	338	14	216
21-40	248	7125	133	3832	31	889	75	2161	12	316	15	466
41-60	72	3555	38	1891	7	348	21	1045	6	308	4	190
61-100	61	4683	30	2293	7	514	15	1177	2	166	6	436
101-200	32	4226	18	2375	1	119	10	1338	6	804	1	114
201-300	6	1439	4	1063	1	258	2	547	1	258	0	0
301-400	1	321	5	1729	0	0	4	1360	1	369	0	0
401-582	0	0	4	1906	2	866	1	458	1	582	0	0
Total	2,249	30,261	1,247	20,810	214	4,274	802	11,594	113	3,303	118	1639

NP=Number of Papers, TC=Total Citations.

Table 9: **Pattern of Altmetric Attention Score.**

AAS Score Range	Closed Access		Open Access		OA Categories							
					Green OA		Gold OA		Hybrid OA		Bronze OA	
	NP	AAS	NP	AAS	NP	AAS	NP	AAS	NP	AAS	NP	AAS
0.25-0.99	162	65.65	120	46.50	25	8.75	79	30.50	7	3.00	9	4.25
1.00-10.99	266	819.31	212	731.72	36	122.36	133	438.74	27	110.39	16	60.23
11.00-30.99	27	465.87	48	868.21	7	153.91	28	469.54	9	170.93	4	73.84
31.00-50.99	1	44.08	19	797.15	3	117.03	10	419.84	5	228.28	1	32
51.00-100.99	3	226.29	6	429.46	1	54.50	4	285.08	0	0.00	1	89.88
101.00-200.99	1	112.50	4	514.96	1	127.51	3	387.45	0	0.00	0	0.00
201.00-1000.99	0	0.00	9	3711.08	1	226.88	5	1735.57	3	1748.63	0	0.00
1001.00-2609.94	1	2609.95	3	3525.33	0	0.00	2	2522.79	0	0.00	1	1002.54
Total	461	4343.65	421	10624.42	74	810.93	264	6289.51	51	2261.23	32	1262.74

NP=Number of Papers, AAS=Altmetric Attention Score.

CC-BY license. Figure 2 visually illustrates the distribution of citations between open and closed access publications.

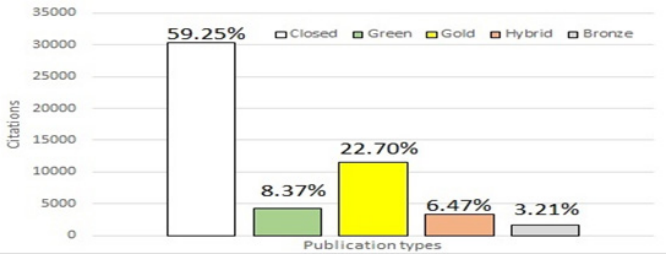

Figure 2: Citations share of closed and open access publications.

Altmetric Attention Score in Open Access Dairy Research

The Altmetric Attention Score (AAS) offers insights into the socio-academic impact of scholarly publications. In this study, the Altmetric Attention Score of Indian dairy research publications was examined, with the AAS for closed access publications specifically detailed in Table 9 for comparison with the AAS of open access publications. Upon analyzing the altmetric response data, it was determined that 24.92% (n=882 out of 3539) of publications had received altmetric attention scores. These 882 publications, encompassing both closed and open access, collectively earned an altmetric attention score of 14,968.06. In the open access category, 421 publications (47.73% of the total 882 responded publications) received a total altmetric attention score of 10,624.42, constituting 70.98% of the total AAS. Figure 3 visually illustrates that Gold OA publications received the highest AAS, followed by Hybrid OA, Bronze OA, and Green OA. Specifically, Gold OA received 42.02% of the total AAS. Figure 3

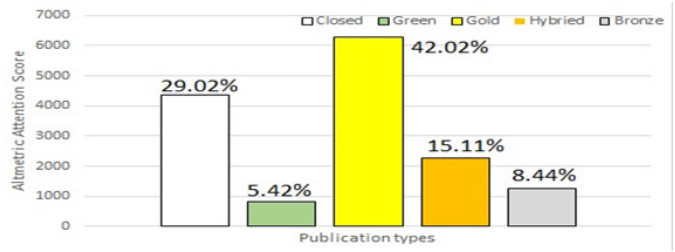

Figure 3: Altmetric Attention Score of Publications.

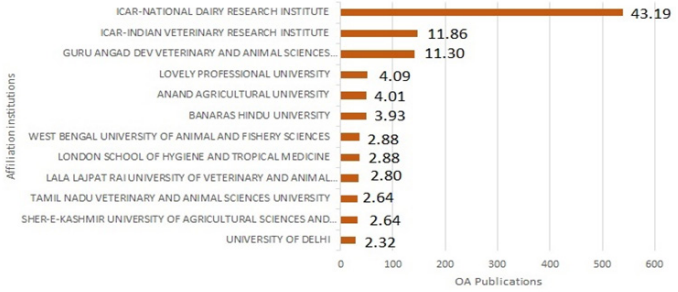

Figure 4: Top ten OA productive institutions.

underscores that open access publications received a higher AAS when compared to closed access publications.

Productive Institutions in Open Access

Figure 4 emphasizes the top ten open access productive institutions. Out of a total of 1,530 institutions contributing to dairy research, ICAR-National Dairy Research Institute emerged as the top contributor of OA publications in Dairy research, with 539 publications (43.19% of OA publications). It was followed by ICAR-Indian Veterinary Research Institute (*n*=148, 11.86%), Guru Angad Dev Veterinary and Animal Sciences University (*n*=141, 11.30%), Lovely Professional University (4.09%), Anand Agricultural University (4.01%), and Banaras Hindu University (3.93%).

Key Findings

The proportion of publications with DOIs increased from 2013 to 2022, while those without DOIs decreased.

Over 30% of publications were open access, with the highest share in 2013.

Licenses were attached to 57.77% of open access publications, with 12 document types identified.

"Article" was the dominant document type, constituting 71.77% of total publications.

The top ten open access journals contributed to 33.97% of open access publications.

Elsevier BV and European Bioinformatics Institute were the top publishers by publication count.

Open access publications received more citations and altmetric attention.

ICAR-National Dairy Research Institute was the top contributor to open access dairy research publications.

CONCLUSION

This study presents a comprehensive analysis of the open access status of dairy research publications in India over the past decade. The research reveals various perspectives, including the growth of open access publications, the utilization of licenses in open access, types of documents, leading journals and publishers in open access, authorship trends, citation patterns, altmetric attention scores for dairy research, and the top-performing Indian institutions in terms of open access. Data were sourced from Scopus for the decade 2013-2022 using a formal search query, and a data carpentry approach was employed to process the retrieved data for analysis. The results indicate a significant number of publications with DOIs, with the ratio of DOI publications increasing over time, while non-DOI publications decreased from the initial days. The study further reveals a moderate increase in closed access publications, but over the past ten years, the ratio of open access publications consistently exceeded 30%. Among all open access routes, Gold OA demonstrated a significant increase in publication counts. The study identified 10 types of open access licenses attached to the OA publications. The majority of OA publications were published with a license, with CC-BY being highly attached. Notably, a significant portion of Green OA publications lacked licensing, while Hybrid OA had the highest proportion of licensed publications. Document types were comprehensively examined, revealing ten distinct types, with "Article" being predominant, followed by "Review" and "Conference Paper."

In terms of leading journals, the study found that 427 journals contributed to 1,248 open access publications and the "Journal of Food Science and Technology" held the highest number of publications in dairy research. The study identified 136 publication agencies, with Elsevier BV leading in publications, citations, and contributing to 27% of citations share. Authorship patterns were examined, revealing multiple authorship patterns as the highest, with the three-authorship pattern being predominant. Gold OA-related publications had the highest number of authors. In terms of citation-related data, open access publications had a higher citation ratio than closed access publications. Hybrid OA had the highest citation ratio, while Gold OA had the highest share of citations. Altmetric Attention Scores for open access publications showed a 70% score, compared to 30% for closed access publications. Finally, the study identified ICAR-National Dairy Research Institute as the top contributor of OA publications in dairy research, with 539 publications. In summary, this study sheds light on the growth of open access publications in Indian dairy research, acknowledging various contributing factors and commending the commendable growth observed during the study period.

Limitations and Future Prospects of the Study

The analysis of the present study relies on data obtained from the Scopus database over a decade, limiting it to a specific period and database. This limitation can be addressed by future researchers who may expand the study by incorporating additional leading databases and extending the time period. A comparative study within the same domain, utilizing two different datasets, could be conducted to gain a deeper understanding of the progress in open access literature.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCES

Abraham, T., Vaidya, N., Kumar, V., Gutam, S., & Guttikonda, A. (2009). Open access Journal Publishing in the Agricultural Sciences. In PKP international scholarly publishing conference (pp. 8–10).

Agrawal, A. N. (1980). Indian Agriculture: Problems, Progress and prospects.

Archambault, É., Amyot, D., Deschamps, P., Nicol, A., Provencher, F., Rebout, L., & Roberge, G. (2014). Proportion of open access papers published in peer-reviewed journals at the European and world Levels: 1996–2013. Digital commons. Retrieved from https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1007

Archambault, E., Amyot, D., Deschamps, P., Nicol, A., Rebout, L., & Roberge, G. (2013). Proportion of open access peer-reviewed papers at the European and world levels—2004-2011. European Commission. Retrieved from https://science-metrix.com/pdf/SM_EC_OA_Availability_2004-2011.pdf

Arunachalam, S. (1998). Agricultural Research in India: A Profile Based on CAB Abstracts. https://repository.arizona.edu/handle/10150/105177 (pp. 1990–1994).

Arunachalam, S., & Umarani, K. (2001). Mapping agricultural research in India: A profile based on CAB Abstracts 1998. Current Science, 81(8), 896-906.

Bhat, M. H. (2009). Open access publishing in Indian premier research institutions. Information Research – An International Electronic Journal, 14(3).

Björk, B. C. (2004). Open access to scientific publications: An analysis of the barriers to change? Information Research, 9(2).

Chakravarty, R., & Mahajan, P. (2011). Open access journals initiatives in India. International Journal of Information Dissemination and Technology, 1(1), 9.

Chan, L., Cuplinskas, D., Eisen, M., Friend, F., Genova, Y., & Guédon, J. C. (2002). Budapest open access initiative. https://www.budapestopenaccessinitiative.org/read/

Fernandez, L. (2006). Open access initiatives in India: An evaluation. Partnership, 1(1), 1–22. DOI: 10.21083/partnership.v1i1.110

Guttikonda, A., & Gutam, S. (2009). Prospects of open access to Indian agricultural research: A case study of ICAR. First Monday, 14(7), 1–11. DOI: 10.5210/fm.v14i7.2488

Keisham, S., & Sophiarani, S. (2008). Open access. Journal and Open Access Initiatives in India. https://ir.inflibnet.ac.in/bitstream/1944/1134/1/19.pdf (pp. 205–216).

Kumar, H. K., Srinivasa, V., Reddy, M. B., Chandra, B., Chandra, B., & Chandra, B. (2012). India's contribution to agriculture and food sciences through open access literature. DESIDOC Journal of Library and Information Technology, 32(1), 53–58. DOI: 10.14429/djlit.32.1.1406

Lal, A., Panda, K. C., & Panda, K. C. (1996). Research in plant pathology: A bibliometric analysis. Library Science with a slant to documentation and information studies, 33(3), 135–147.

Maddi, A., Lardreau, E., & Sapinho, D. (2021). Open access in Europe: A national and regional comparison. Scientometrics, 126(4), 3131–3152. DOI: 10.1007/s11192-021-03887-1

Martín-Martín, A., Costas, R., van Leeuwen, T., & Delgado López-Cózar, E. (2018). Evidence of open access of scientific publications in Google Scholar: A large-scale analysis. Journal of Informetrics, 12(3), 819–841. DOI: 10.1016/j.joi.2018.06.012

Mukhopadhyay, P. (2022). How green is my Valley? Measuring open access friendliness of Indian Institutes of Technology (IITs) through data carpentry. https://zenodo.org/record/6511080. In Panorama of open access: Progress, practices and prospects, Ess Ess (pp. 67–89).

Nazim, M. (2021). Analysing open access uptake by academic and research institutions in India. DESIDOC Journal of Library and Information Technology, 41(2), 108–115. DOI: 10.14429/djlit.41.02.16324

Piwowar, H., Priem, J., & Orr, R. (2019). The Future of OA: A large-scale analysis projecting Open Access publication and readership. bioRxiv. DOI: 10.1101/795310/1

Piwowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B., Farley, A., West, J., & Haustein, S. (2018). The state of OA: A large-scale analysis of the prevalence and impact of Open Access articles. PeerJ, 6, e4375. DOI: 10.7717/peerj.4375

Priem, J., Piwowar, H., & Orr, R. (2022). OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. arXiv preprint arXiv:220501833.

Raychoudhury, N., & Sarkhel, J. K. (2011). Mapping Agricultural Research in West Bengal 1993–2007: A bibliometric study. COLLNET Journal of Scientometrics and Information Management, 5(1), 99–113. DOI: 10.1080/09737766.2011.10700905

Robinson-Garcia, N., Costas, R., & van Leeuwen, T. N. (2020). Open Access uptake by universities worldwide. PeerJ, 8, e9410. DOI: 10.7717/peerj.9410

Robinson-Garcia, N., Costas, R., & van Leeuwen, T. N. Indicators of Open Access for universities, arXiv,2019. DOI: 10.48550/arXiv.1906.03840

Roy, A., & Mukhopadhyay, P. (2022). Assessing open access friendliness of National Institutes of Technology (NITs) A data carpentry approach. DESIDOC Journal of Library and Information Technology, 42(5), 331–338. DOI: 10.14429/djlit.42.5.18263

Roy, A., & Mukhopadhyay, P. (2022). Measuring open access friendliness of Indian central universities through data carpentry. SRELS Journal of Information Management, 59(3), 131–139. DOI: 10.17821/srels/2022/v59i3/17010

Roy, A., & Mukhopadhyay, P. (2022). Measuring the open access friendliness of Indian institutions through data carpentry, CALIBER. https://ir.inflibnet.ac.in/bitstream/1944/2398/1/41.pdf (pp. 498–508).

Roy, A., & Mukhopadhyay, P. (2022). Measuring the open access friendliness of the state universities in India through data carpentry. Annals of Library and Information Studies, 69(3), 225–237. DOI: 10.56042/alis.v69i3.63837

Sagar, A., Kademani, B. S., & Bhanumurthy, K. (2014). Agriculture research in India: A scientometric mapping of publications. DESIDOC Journal of Library and Information Technology, 34(3), 206–222. DOI: 10.14429/djlit.34.5022

Sangam, S. L., Arali, U. B., Patil, C. G., & Gani, S. R. (2014). SR Research trends in genetics: Scientometric profile of selected Asian countries. DESIDOC Journal of Library and Information Technology, 34(3), 248–256. DOI: 10.14429/djlit.34.5802

Shukla, P., & Singh, A. P. (2009). Open access initiatives for agricultural information transfer systems in India. Agricultural Engineering, 7(6).

Singh, V. K., Piryani, R., & Srichandan, S. S. (2020). The case of significant variations in gold–green and black open access: Evidence from Indian research output. Scientometrics, 124(1), 515–531. DOI: 10.1007/s11192-020-03472-y

Suber, P. (2012). Open access. The MIT Press.

Suryanarayana, Y. V. (2002). Tobacco research publications: Global scenario. SRELS Journal of Information Management, 39(2), 183–194.

Tripathi, H. K., Raj, H., Kumar, S., & Mapping of Research Output of Animal Science Division. (2013). Library Herald, 51(1), 50–65.

Cite this article: Roy A, Shukla A. Unveiling the Status of Open Access Dairy Research in India through Data Carpentry. Journal of Data Science, Informetrics, and Citation Studies. 2024;3(1):1-10.