Journal of Scientometric Res. 2020; 9(2):96-110
http://www.jscires.org

Research Article

The Status and Patterns of Open Access in Research Output of Most Productive Indian Institutions

Satya Swarup Srichandan[1], Rajesh Piryani[2], Vivek Kumar Singh[1,*], Sujit Bhattacharya[3]

[1]Department of Computer Science, Banaras Hindu University, Varanasi, Uttar Pradesh, INDIA.
[2]Department of Computer Science, South Asian University, New Delhi, INDIA.
[3]CSIR-NISTADS and AcSIR NISTADS Campus, New Delhi, INDIA.

ABSTRACT
Open Access is emerging as an important movement worldwide since last few years, triggered mainly by the high subscription cost of pay walled journals that create barriers in universal dissemination of knowledge reported in those journals. The paywall barriers to access of knowledge has become so problematic that even institutions in the developed countries are not only cancelling subscriptions but also mandating it for their researchers to either publish in open access journals or at least deposit their research papers in Institutional Repositories. The high subscription cost of journals is a more serious issue for developing countries, as it takes away institutional resources that can be used for other productive purposes. India has taken several steps in promoting open access, including release of an open access policy by Ministry of Science and Technology, however, it is not very clear that how effective these initiatives have been. This paper intends to address this issue. It examines published output, indexed in Web of Science, from 100 most productive institutions in India and analyze how much research output coming from them are available in Open Access (OA). The paper further analyzes availability of research papers from these institutions in the popular pirate site Sci-Hub. It is interesting to observe that legal OA percentages are significantly lesser than the Sci-Hub availability for all the institutions, an indication that the existing systems for promoting open access in India are not working efficiently. At the end, the paper also presents statistics about number of papers deposited in three central institutional repositories in India. These statistics provides an indication of the extent to which these repositories have been able to promote open access in India. The paper concludes by pointing to some factors that impede Open Access in India.

Keywords: Open Access, Sci-Hub, Unpaywall, Gold Open Access, Green Open Access, Black Open Access, Institutional Repository.

Correspondence
Vivek Kumar Singh
Department of Computer Science, Banaras Hindu University, Varanasi-221005, Uttar Pradesh, INDIA.
Email: vivek@bhu.ac.in

Received: 10-11-2019
Revised: 13-02-2020
Accepted: 15-06-2020
DOI: 10.5530/jscires.9.2.13

INTRODUCTION

The increasing subscription cost of journals from large publishing houses is becoming a severe deterrent in universal open access to knowledge. Not only institutions in developing countries are finding it difficult to pay these costs but several institutions in the developed world are also cancelling subscriptions of many journals. Many top Universities in the world are now asking their researchers to publish in open access journals so as to make the knowledge generated accessible, without barriers, to the world. Institutions are creating Institutional Repositories (IRs) to keep copies of papers published by their researchers in different journals. The top funding research organizations of the world are now making it mandatory for their researchers to submit pre- or post-print of their research papers in different repositories, either institutional or disciplinary. Organizations like US National Institutes of Health, US National Science Foundation, Welcome Trust, Bill and Melinda Gates Foundation, the European Commission etc. are some examples. Further, almost all major publishing house are allowing submission of papers that have been earlier submitted as pre-print in repositories.

India has also tried to move towards Open Access, both through efforts at the national level as well as at the institutional level. Open Access culture in India was initiated by Indian physicists back in the early 90s when they started depositing their preprints in arXiv. Later, they were joined by Mathematicians, Computer Scientists, Biologists etc. Since then there have been several initiatives taken, though the

impact was not as expected. Some of these early initiatives include Mirror server for arXiv set up by Institute of Mathematical Sciences (IMSc) at Chennai, the Vidyanidhi Digital Library for electronic thesis and dissertations from the university of Mysore in 2002[1] and EPrints@IISc electronic repository by Indian Institute of Science, Bangalore in 2002.[2] The open access and open educational resources were discussed by the National Knowledge Commission in a report in 2007.[3] In 2009, CSIR headquarters sent a memorandum to all of its 38 laboratories in the country to set up institutional open access repositories, as detailed ahead. The work of Das[4] and Arunachalam and Muthu[5] can be referred to know about initial open access initiatives and projects in India.

The Government of India's most recent effort in this direction was the release of the Department of Biotechnology (DBT) and Department of Science and Technology (DST) Open Access Policy in 2014.[6] The DBT and DST, the two main research Departments of the Ministry of Science and Technology of Government of India have jointly issued the policy. It states that "*since all funds disbursed by the DBT and DST are public funds, it is important that the information and knowledge generated through the use of these funds are made publicly available as soon as possible....*". The fundamental guiding principal of this policy is the fact that public funded research outcomes should be publicly available. The policy also envisaged that open access would allow percolation of cutting-edge research in Higher Education curricula, which in turn will raise the standards of technical and scientific education in the country. Through the policy, institutions were encouraged to setup institutional repositories (IRs) and to deposit all their research outcome in them. A central harvester *sciencecentral.in* was also created and it was expected that all institutional repositories would eventually link to it. All institutions receiving funds from DBT and DST are expected to mandatorily follow the policy.

In another similar step, the open access mandate[1] has been issued by Council of Scientific and Industrial Research (CSIR) which instructs the setup of interoperable institutional open access repositories at all CSIR laboratories. Through this mandate CSIR took the initiative to "*lead the open access movement within the country*". The *csircentral.net* established by CSIR-URDIP is the central harvester that is expected to link different institutional repositories created by individual CSIR labs. The Indian Council of Agriculture Research (ICAR) has also created an open access policy,[2] which required setting up of open access institutional repository by each ICAR institute. It has also setup a one stop access platform, to provide access to all the agricultural knowledge generated in ICAR, known as *Krishikosh*.[3] In 2018, the Delhi Declaration on Open Access[4] was signed by a group of academician and open access enthusiasts, which advocated for "the practice of open science" and "adoption of open technologies for the development of models for sharing science and scholarship". The IndiaRxiv[5] of the Open Science project is the most recent addition to the open access cause. Open Access India has initially launched AgriXiv, a preprints repository for agriculture and allied sciences and then launched IndiaRxiv. However, this preprints repository has only 67 preprints available as on 16th Feb 2020, since its launch in April 2019 and is yet to pick its momentum.

Despite all these initiatives and efforts, India has somehow lagged behind in ensuring open access to its research output. There are no proper estimation of the research emerging from Indian institutions which is in open access. It is in this context that this paper tries to measure the current status of open access in research output from 100 most productive institutions in India. These institutions provide a good estimation of research available in Indian institutions in open access as they together account for about 82% of the total research output from India for the year 2016 as indexed in Web of Science. Papers covered in WoS are regarded as a benchmark of quality by several institutions and there is a general tendency for researchers to access papers that are in this database. To provide another perspective and complement this study, a detailed examination of volume of research papers in various institutional repositories (IR) in India is also examined.

In addition to computing open access levels of research output indexed in Web of Science and in institutional repositories in India, the study also present statistics of availability of these papers in Sci-Hub. Sci-Hub provides a questionable route for access to scientific papers as it ignores copyright clauses to allow users to download papers which are not openly licensed content. Our motive at examining paper availability in SCI-Hub is not an endorsement of this type of database but to underscore that lack of Open Access may lead to these types of behaviors. Apart from this it can lead to legal consequences for readers, authors and institutions to which they are affiliated, if they use this approach.

The paper focusses on three questions:

First, how much scholarly research output from the 100 most productive Indian Institutions is available in legal (such as gold, green and bronze) and black access models?

Second, how much research output is available in open access through IRs in India?

Third, what are the plausible factors that impede open access in India?

Open Access Types

There exist multiple forms of open access. Some journals are absolutely free and completely open access, where articles are

1 http://www.csircentral.net/mandate.pdf
2 https://icar.org.in/hi/node/5542

freely available to everyone. Then, there are paywalled journals that requires reader to pay before accessing the article. But these journals may also make some articles freely available to everyone, either after payment of an article processing charge or after a particular period from publication date. Open access articles are classified into following main categories based on the levels and types of open access:

Gold open access: This refers to an article which is freely available for everyone as it is published in a journal that is open access. Generally, the Directory of Open access Journals (DOAJ) includes qualitative open access journals, after a self-nomination and evaluation process. In this type of open access, copyright is usually retained by the authors.

Green open access: This refers to articles which are accessible free of cost as they are uploaded on some repository (either institutional or disciplinary), though they may be published in pay walled journals. However, the reuse rights are restricted in such case and in most of the cases, the articles can be uploaded on such a repository after a certain period of time from their publication in the journal (varying from 6 to 48 months).

Hybrid open access: This refers to an article which becomes open access after an article processing charge is paid (along with some agreement with the journal) by the authors or their institution. Such articles are published in closed journals.

Bronze open access: This refers to an article which is freely available to read but without a license.

Black open access: This refers to an article that is shared on illegal pirate sites, such as Sci-Hub[6] or LibGen.[7] However, this type is not well recognized as open access in the literature.

Closed access: This refers to all other articles that are not openly accessible in legal forms. The copyright is with the publisher and readers need to pay to access the paper.

Related Work

There exist several previous studies that tried to understand and characterize open access (OA) patterns in research outputs at the international level. Hajjem *et al*.[7] is one of the earliest studies to have analyzed open access availability of articles and found that OA articles have comparatively more citations than non-OA ones. Bjork *et al*.[8-13] through their multiple studies during 2010 to 2017, analyzed the open access patterns in scientific publishing, with varied data. Archambault *et al*.[14,15] analyzed the proportion of open access peer-reviewed papers at European and world levels for 2004-2011 and 1996-2013 time periods, respectively. They have shown that several countries, including Brazil, Switzerland, Netherlands, US have more than 50% of the research articles freely available. Piwowar *et al*.[16] used three different samples of 100K articles each drawn from Web of Science (WoS), CrossRef and Unpaywall for publications from worldwide and found that about 28% the scholarly articles are available in open access. Bosman and Kramer[17] collected data from Web of Science using its oaDOI service and explored open access levels across research fields, languages, countries, institutions, funders and topics and found high variations in open access levels on all these dimensions.

There are, however, very few recent studies on open access levels in publications from India. Among the recent studies, Kumar and Mahesh[18] tried to analyze the Institutional Repositories in India, as a means of providing open access to articles. They analyzed the statistics of submission of papers for about one year and found that submissions to the Institutional Repositories were very low, with some not even getting a single paper deposited in the whole year. They concluded that the Institutional Repositories in India have not really picked up in terms of papers deposited. Another recent article in Nature[19] talked about the newly proposed IndiaRxiv repository, with the caution that the performance of Institutional Repositories in India is not very good. Piryani *et al*.[20] is another recent study that focused on measuring open access levels for India as a country. They tried to analyze the overall level of open access in Indian research output by taking data from the Web of Science for all publications during 2016. They conclude that the overall open access level in Indian research output is about 24% of the total output, which is less than the world average. Another study by Singh *et al*.[21] have analyzed Indian research output data in Web of Science for the period 2014-18 and obtained OA evidence from Unpaywall as well as volume of papers available for free download from Sci-Hub. However, both of these previous studies only looked at overall data for India as a county and did not go at the institutional level. Therefore, there is no evidence available about what proportion of research papers from different institutions in India are available as OA. Further, no recent analysis is available about number of papers available in Indian IRs.

Data and Methodology

The 100 most productive Indian institutions (including institution systems) account for a total of 62,688 publications out of total 76,530 publications indexed in Web of Science for India for the publication year 2016. This is about 82% of the total research output from India for the year 2016 as indexed in Web of Science. All these 62,688 publication records are then scanned one by one to find out if they are available in open access in some platform.

Thus, the data for analysis in this work is obtained from four sources: (a) Web of Science[3] (WoS) database, (b) Unpaywall[9] portal, (c) Sci-Hub and (d) main IRs in India.

For all the publication records downloaded from WoS, the OA evidence is obtained from the Unpaywall portal through API calls. The data downloading from WoS and OA evidence from the Unpaywall portal was obtained in the month of July 2019.

The publication record data from WoS was used to make an automated lookup in Sci-Hub website. The crawling was done using a custom Python script to identify which of the DOIs in our WoS data have full text available in Sci-Hub. Thus, for all the 62,688 publication records obtained from the WoS database, the Sci-Hub portal was queried and evidence of availability of full text was recorded.

Further, as was observed by Das and Dutta,[22] there were 726 open access journals covered by WoS in 2017, which covers 7% of the total open access articles globally, therefore, it is clear that capturing open access output from WoS only would limit any study. It is for this reason, we also tried to look at data in IRs in India as this data complements the study and provides a more informed assessment of institutional research that is in open access from India. Data from the three main central portals and their associated IRs, as explained later, was obtained and analyzed.

The publication records data has been analyzed computationally by writing programs in Python and R. Standard computational methods are used for computing results and generating plots. As stated earlier, the OA evidence was obtained from Unpaywall through an API lookup in the portal. One important point of analysis in the paper is to look at disciplinary variations in open access. For this purpose, each publication record is tagged into one of the 14 broad research disciplines, as proposed in a previous work.[23] The Web of Science Category (WC) field is used for this tagging. These 14 research disciplines are: Agriculture (AGR), Art and Humanities (AH), Biology (BIO), Chemistry (CHEM), Engineering (ENG), Environment Science (ENV), Geology (GEO), Information Sciences (INF), Material Science (MAR), Mathematics (MAT), Medical Science (MED), Multidisciplinary (MUL), Physics (PHY) and Social Science (SS). The analytical results are presented in different tables and figures.

Open Access Evidence from Unpaywall

The 100 Indian institutions covered in this study account for 62,688 publication records. Out of 62,688 publication records, only 14,454 records are in Unpaywall, i.e. found to be available in some form of open access. In percentage terms, only about 23% of the combined output of 100 most productive institutions in India is available in open access.[10] The number of open access articles from all these institutions combined contributes about 83.65% of the total number of open access articles for India for the year 2016. When we look at open access categories, we observe that out of the combined output from the 100 institutions, 5,893 records are Gold open access (~41% of total open access articles), 3,844 records are Green open access (~27% of total open access articles) and 2,580 records are Bronze open access (~18% of total open access articles). Thus, the Gold open access type is the most prevalent type of open access.

We tried to see if the open access levels in the combined output of 100 institutions are similar for all disciplines or there are variations in different disciplines. Figure 1 the disciplinary distribution of open access articles in the combined output of the 100 institutions. It can be seen that PHY discipline accounts for the highest proportion (23%) of articles in the total open access articles for all the institutions taken together. This is followed by MED discipline with a 17% share and MUL discipline with a 13% share. Thus, PHY, MED and MUL taken together contribute more than 50% of open access articles in the combined output of the 100 institutions. SS, MAT and (surprisingly) INF disciplines have a very low contribution to open access articles, possibly also because they have less volume of output too. PHY discipline is the most interesting case with only 12.25% contribution in the combined output of the 100 institutions but when it comes to contribution to open access articles, it is much higher at 23%. One possible reason for this could be the existence of the well-known arXiv repository where Physicists are the key contributors.

After looking at the OA levels for combined data of all institutions, we analyzed OA levels at the level of granularity of individual institutions. Table 1 presents the detailed data for all the 100 institutions (including institution systems). It shows the name and location of the institution, number of records it has in Web of Science, number of records that are available in open access and also percentage contribution that this institution makes to total output as well as to total open access articles from India. In order to understand open access levels in each of the institutions in more detail, we categorize institutions in three categories: OA_Low, OA_Med and OA_High corresponding to open access percentage levels of below 25%, between 25 to 45% and above 45%, respectively. It can be observed that out of 100 Institutions, there are 58 Institutions under OA_Low category, 32 Institutions under OA_Med category and 10 under OA_High category. The statistics for these 100 institutions are also plotted in Figure 2 as a scatter plot for an alternative way of visualization and understanding. Here, the x-axis denotes the number of papers

3 https://www.webofknowledge.com/

Table 1: Open Access (OA) Levels in 100 most productive Indian Institutions.

S. No.	Institution Name	No. of articles in WoS	Articles with DOI	No. of articles that are OA	OA articles as Percentage of total articles[1]	Proportionate Contribution to the total OA of India[2]
	Institution Systems/ Govt Departments					
1	Council of Scientific Industrial Research (CSIR)	8041	5027	965	19.20	5.58
2	Indian Council of Agricultural Research (ICAR)	3774	1982	588	29.67	3.40
3	Department of Science Technology India (DST)	2146	2053	719	35.02	4.16
4	Department of Biotechnology (DBT)	890	825	464	56.24	2.69
5	Defence Research Development Organisation (DRDO)	856	821	161	19.61	0.93
6	Ministry of Earth Sciences (MOES) India	414	356	96	26.97	0.56
7	Indian Space Research Organisation (ISRO)	310	289	63	21.80	0.36
8	Indian Council of Medical Research (ICMR), New Delhi	257	235	108	45.96	0.63
	Individual Institutions					
1	Indian Institute of Science, Bangalore (IISc)	1861	1805	529	29.31	3.06
2	Indian Institute of Technology, Kharagpur	1666	1589	206	12.96	1.19
3	All India Institute of Medical Sciences (AIIMS), New Delhi	1621	1217	474	38.95	2.74
4	Bhabha Atomic Research Center (BARC)	1491	1464	292	19.95	1.69
5	Indian Institute of Technology, Delhi	1479	1423	223	15.67	1.29
6	Indian Institute of Technology, Bombay	1412	1371	288	21.01	1.67
7	Indian Institute of Technology, Madras	1393	1363	263	19.30	1.52
8	University of Delhi (DU)	1226	1141	291	25.50	1.68
9	Banaras Hindu University, Varanasi (BHU)	1210	1095	251	22.92	1.45
10	PGIMER Chandigarh	1197	942	340	36.09	1.97
11	Indian Institute of Technology, Roorkee	1190	1162	134	11.53	0.78
12	Indian Institute of Technology, Kanpur	1088	1057	246	23.27	1.42
13	Anna University, Chennai	1023	912	156	17.11	0.90
14	Jadavpur University, Kolkata	972	923	120	13.00	0.69
15	Tata Institute of Fundamental Research (TIFR)	955	912	663	72.70	3.84
16	Vellore Institute of Technology (VIT)	922	854	150	17.56	0.87
17	Indian Institute of Technology, Guwahati	888	853	134	15.71	0.78
18	University of Calcutta, Kolkata	790	744	160	21.51	0.93
19	Manipal University, Karnataka	786	687	283	41.19	1.64
20	Indian Institute of Technology (ISM Dhanbad)	746	713	66	9.26	0.38
21	Panjab University, Chandigarh	682	337	99	29.38	0.57
22	Aligarh Muslim University (AMU)	625	585	168	28.72	0.97
23	University of Hyderabad, Hyderabad	613	595	169	28.40	0.98
24	Savitribai Phule Pune University, Pune	600	569	183	32.16	1.06
25	Jawaharlal Nehru University (JNU), New Delhi	596	545	175	32.11	1.01
26	National Institute of Technology, Rourkela	550	538	71	13.20	0.41
27	Birla Institute of Technology and Science, Pilani	497	473	104	21.99	0.60

28	Indian Institute of Technology (IITBHU) Varanasi	491	475	50	10.53	0.29
29	Tata Memorial Hospital (TMH)	483	291	105	36.08	0.61
30	Saha Institute of Nuclear Physics, Kolkata	468	466	278	59.66	1.61
31	Thapar University	465	433	58	13.39	0.34
32	Christian Medical College Hospital (CMCH), Vellore	465	376	194	51.60	1.12
33	Bharathiar University, Coimbatore	462	425	63	14.82	0.36
34	Annamalai University, Tamil Nadu	444	397	76	19.14	0.44
35	Indian Association for the Cultivation of Science (IACS), Jadavpur	430	422	84	19.91	0.49
36	Institute of Chemical Technology (ICT), Mumbai	403	398	37	9.30	0.21
37	National Institute of Mental Health Neurosciences India	399	330	120	36.36	0.69
38	Amity University, NOIDA	399	371	84	22.64	0.49
39	Indian Institute of Science Education Research (IISER), Pune	396	384	200	52.08	1.16
40	Indian Statistical Institute (ISI) Kolkata	395	375	149	39.73	0.86
41	Sanjay Gandhi Postgraduate Institute of Medical Sciences (SGPGI), Lucknow	394	302	102	33.77	0.59
42	Indira Gandhi Centre for Atomic Research (IGCAR), Tamil Nadu	391	386	43	11.14	0.25
43	National Institute Technology (NIT), Tiruchirappalli	387	378	37	9.79	0.21
44	Guru Nanak Dev University, Amritsar	387	377	57	15.12	0.33
45	Jamia Millia Islamia (JMI), Delhi	382	366	82	22.40	0.47
46	National Institute of Pharmaceutical Education Research (NIPER), Mohali	372	352	82	23.30	0.47
47	Visva Bharati University, Burdwan	355	345	146	42.32	0.84
48	Indian Institute of Engineering Science Technology (IIEST), Shibpur	341	331	35	10.57	0.20
49	Shanmugha Arts Science Technology Research Academy (SASTRA), Tamil Nadu	332	314	81	25.80	0.47
50	Indian Institute of Technology, Indore	331	330	98	29.70	0.57
51	Jawaharlal Nehru Center for Advanced Scientific Research (JNCASR)	326	321	100	31.15	0.58
52	Indian Institute of Science Education Research (IISER), Kolkata	320	318	136	42.77	0.79
53	Bharathidasan University, Tamil Nadu	317	300	83	27.67	0.48
54	University of Madras, Chennai	317	301	54	17.94	0.31
55	Osmania University, Telangana	313	289	47	16.26	0.27
56	National Institute of Science Education Research (NISER)	307	306	194	63.40	1.12
57	Sri Venkateswara University, Andhra Pradesh	300	277	59	21.30	0.34
58	Jamia Hamdard University, New Delhi	294	275	70	25.45	0.41
59	Tezpur University, Assam	293	288	42	14.58	0.24
60	Pondicherry University, Pondicherry	292	273	52	19.05	0.30
61	Indian Institute of Technology, Hyderabad	286	280	68	24.29	0.39
62	Shivaji University, Maharasthra	285	275	29	10.55	0.17

63	Amrita Vishwa Vidyapeetham University, Tamil Nadu	284	247	60	24.29	0.35
64	King Georges Medical University (KGMU), Lucknow	272	223	78	34.98	0.45
65	Indian Institute of Science Education Research (IISER), Bhopal	261	257	125	48.64	0.72
66	Public Health Foundation of India	257	245	185	75.51	1.07
67	Visvesvaraya National Institute of Technology (VNIT), Nagpur	247	236	32	13.56	0.19
68	Birla Institute of Technology Mesra	240	229	36	15.72	0.21
69	Physical Research Laboratory (PRL), India	239	224	147	65.63	0.85
70	University of Allahabad	232	201	31	15.42	0.18
71	Madurai Kamaraj University, Tamil Nadu	228	218	31	14.22	0.18
72	Homi Bhabha National Institute, Mumbai	228	228	60	26.32	0.35
73	Jawaharlal Institute of Postgraduate Medical Education Research (JIPMER)	225	205	81	39.51	0.47
74	University of Mysore	225	208	57	27.40	0.33
75	University of Kashmir	224	206	45	21.84	0.26
76	Andhra University	222	198	52	26.26	0.30
77	National Institute of Technology (NIT), Durgapur	220	214	24	11.21	0.14
78	Alagappa University, Tamil Nadu	218	210	16	7.62	0.09
79	Maharaja Sayajirao University Baroda	217	212	41	19.34	0.24
80	Sardar Vallabhbhai National Institute of Technology (SVNIT)	217	197	30	15.23	0.17
81	Cochin University Science Technology (CUST), Kerala	217	196	35	17.86	0.20
82	PUNJAB AGRICULTURAL UNIVERSITY	213	162	41	25.31	0.24
83	Kalyani University, West Bengal	213	201	29	14.43	0.17
84	National Institute of Technology (NIT), Karnataka	213	207	20	9.66	0.12
85	BOSE Institute	212	208	107	51.44	0.62
86	Jaypee University Of Information Technology	212	208	42	20.19	0.24
87	Institute for Plasma Research IPR	209	209	52	24.88	0.30
88	Central Salt Marine Chemical Research Institute India	207	198	21	10.61	0.12
89	Lucknow University	205	190	16	8.42	0.09
90	Raja Ramanna Centre For Advanced Technology	202	200	49	24.50	0.28
91	L V Prasad Eye Institute	201	169	68	40.24	0.39
92	Inter-University Accelerator Centre	201	199	12	6.03	0.07

[1]Percentage is calculated with respect to number of articles in WoS having a DOI, i.e. column 4.

[2]Percentage calculated with respect to total OA count of India (=17,280)

Note: The 100 institutions include 8 institution systems and 92 individual institutions

indexed in Web of Science from an institution and the y-axis denotes the number of articles that are open access.

Some of the institutions that have higher open access percentages are TIFR at Mumbai, Saha Institute of Nuclear Physics at Kolkata, NISER in Bhubaneswar, PRL in Ahmedabad and PHFI in Delhi. Out of these, most institutions have their main focus on the PHY research area, which has a well-established practice of archiving. DST (along with its associated organizations), AIIMS at New Delhi, PGIMER at Chandigarh, IISER at Kolkata, JNU at New Delhi, IISER at Bhopal are some of the institutions in OA_Med category. These institutions are either in the MED area or are multidisciplinary Universities. The majority of the institutions in OA_Low category are mainly specializing in Engineering and Technology disciplines, such as IIT Kharagpur, IIT Delhi, IIT Roorkee, VIT Vellore, IIT Guwahati, IIT(ISM) Dhanbad, NIT Rourkela etc. Thus, it appears that Engineering and Technology focused institutions, in general, have lower proportion of their research papers available in open access. CSIR has 18.17% of articles in open access, which is reasonable taking into account that it is a conglomerate of 38 laboratories that have disciplinary diversity covering all the disciplines. Thus, the variation in open access across disciplinary boundaries as seen from the overall statistics also has a high probability of getting reflected when aggregation of CSIR output is taken into account. In absolute terms, CSIR

is observed as the largest contributor to open access articles from India.

The research discipline is observed to be an important correlate for a higher proportion of open access articles, as it is seen in institutions that work mainly in the area of Physics. Many institutions working in the area of MED and MUL also have relatively higher proportion of articles available in open access. On the other hand, institutions mainly specializing in Engineering and technology areas, in general, have lower proportion of their research output available in open access. The geographical location of an institution is not found to be a differentiating factor for low or high open access. This could be observed by looking at the fact that some institutions located in big cities like Mumbai and Delhi are found to have low open access levels, on the other hand, some institutions located in small towns are found to have higher open access levels.

Papers available in Sci-Hub

We also looked at how many papers from each institution are available as full text for free download in the popular pirate site Sci-Hub. Our argument is that Indian papers if available in legal forms of open access will not promote illegal access, as researchers, more so in developing countries, are constrained by research funding and hence tend to exploit this type of resources (see for example Gresheke).[24] At the same time, it is equally interesting to note that Sci-Hub is also frequently used for paper downloads from both, the developed and developing countries.

Table 2 shows the number and percentages of papers for each institution available for free full-text download from Sci-Hub. It can be observed that, unlike low levels of legal open access forms, here the availability of articles for free download is much higher for all the institutions. The highest value being 99% articles for Institute of Plasma Research and the lowest being 73.5% articles for Tata Memorial Hospital. Majority of the institutions in the 100-institution set, have more than 90% of their papers available for free download from Sci-Hub. This is a huge contrast with the results of legal OA availability percentages. These values show that something is wrong with the legal forms of open access models as they relate to Indian research. Sci-Hub seems to be complementing the low legal OA availability with, on an average, more than 90% of the articles available for free download. Some studies (for example. Singh *et al.*)[21] Have also found that Sci-Hub seem to complement the availability in different disciplines too, with disciplines having low legal OA levels being preferentially covered in Sci-Hub.

Table 2: Availability of articles from 100 most productive Indian Institutions in Sci-Hub.

S. No.	Institution Name	No. of articles in WoS	Articles with DOI	No. of articles that are Available on Sci-hub	Sci-hub articles as Percentage of total articles[1]	Proportionate Contribution to the total Sci-hub of India[2]
	Institution Systems/ Govt Departments					
1	Council of Scientific Industrial Research (CSIR)	8041	5027	4831	96.10	11.52
2	Indian Council of Agricultural Research (ICAR)	3774	1982	1611	81.28	3.84
3	Department of Science Technology India (DST)	2146	2053	1957	95.32	4.67
4	Department of Biotechnology (DBT)	890	825	777	94.18	1.85
5	Defence Research Development Organisation (DRDO)	856	821	744	90.62	1.77
6	Ministry of Earth Sciences (MOES) India	414	356	337	94.66	0.80
7	Indian Space Research Organisation (ISRO)	310	289	266	92.04	0.63
8	Indian Council of Medical Research (ICMR), New Delhi	257	235	202	85.96	0.48
	Individual Institutions					
1	Indian Institute of Science, Bangalore (IISc)	1861	1805	1730	95.84	4.13
2	Indian Institute of Technology, Kharagpur	1666	1589	1553	97.73	3.70
3	All India Institute of Medical Sciences (AIIMS), New Delhi	1621	1217	977	80.28	2.33
4	Bhabha Atomic Research Center (BARC)	1491	1464	1421	97.06	3.39
5	Indian Institute of Technology, Delhi	1479	1423	1384	97.26	3.30
6	Indian Institute of Technology, Bombay	1412	1371	1323	96.50	3.16
7	Indian Institute of Technology, Madras	1393	1363	1308	95.96	3.12
8	University of Delhi (DU)	1226	1141	1071	93.87	2.55
9	Banaras Hindu University, Varanasi (BHU)	1210	1095	1034	94.43	2.47
10	PGIMER Chandigarh	1197	942	778	82.59	1.86
11	Indian Institute of Technology, Roorkee	1190	1162	1127	96.99	2.69
12	Indian Institute of Technology, Kanpur	1088	1057	1031	97.54	2.46

13	Anna University, Chennai	1023	912	829	90.90	1.98
14	Jadavpur University, Kolkata	972	923	880	95.34	2.10
15	Tata Institute of Fundamental Research (TIFR)	955	912	864	94.74	2.06
16	Vellore Institute of Technology (VIT)	922	854	811	94.96	1.93
17	Indian Institute of Technology, Guwahati	888	853	830	97.30	1.98
18	University of Calcutta, Kolkata	790	744	712	95.70	1.70
19	Manipal University, Karnataka	786	687	605	88.06	1.44
20	Indian Institute of Technology (ISM Dhanbad)	746	713	686	96.21	1.64
21	Panjab University, Chandigarh	682	337	312	92.58	0.74
22	Aligarh Muslim University (AMU)	625	585	546	93.33	1.30
23	University of Hyderabad, Hyderabad	613	595	568	95.46	1.35
24	Savitribai Phule Pune University, Pune	600	569	545	95.78	1.30
25	Jawaharlal Nehru University (JNU), New Delhi	596	545	530	97.25	1.26
26	National Institute of Technology, Rourkela	550	538	525	97.58	1.25
27	Birla Institute of Technology and Science, Pilani	497	473	451	95.35	1.08
28	Indian Institute of Technology (IITBHU) Varanasi	491	475	459	96.63	1.09
29	Tata Memorial Hospital (TMH)	483	291	214	73.54	0.51
30	Saha Institute of Nuclear Physics, Kolkata	468	466	456	97.85	1.09
31	Thapar University	465	433	400	92.38	0.95
32	Christian Medical College Hospital (CMCH), Vellore	465	376	316	84.04	0.75
33	Bharathiar University, Coimbatore	462	425	403	94.82	0.96
34	Annamalai University, Tamil Nadu	444	397	371	93.45	0.88
35	Indian Association for the Cultivation of Science (IACS), Jadavpur	430	422	414	98.10	0.99
36	Institute of Chemical Technology (ICT), Mumbai	403	398	385	96.73	0.92
37	National Institute of Mental Health Neurosciences India	399	330	250	75.76	0.60
38	Amity University, NOIDA	399	371	354	95.42	0.84
39	Indian Institute of Science Education Research (IISER), Pune	396	384	376	97.92	0.90
40	Indian Statistical Institute (ISI) Kolkata	395	375	359	95.73	0.86
41	Sanjay Gandhi Postgraduate Institute of Medical Sciences (SGPGI), Lucknow	394	302	247	81.79	0.59
42	Indira Gandhi Centre for Atomic Research (IGCAR), Tamil Nadu	391	386	381	98.70	0.91
43	National Institute Technology (NIT), Tiruchirappalli	387	378	359	94.97	0.86
44	Guru Nanak Dev University, Amritsar	387	377	361	95.76	0.86
45	Jamia Millia Islamia (JMI), Delhi	382	366	347	94.81	0.83
46	National Institute of Pharmaceutical Education Research (NIPER), Mohali	372	352	327	92.90	0.78
47	Visva Bharati University, Burdwan	355	345	334	96.81	0.80
48	Indian Institute of Engineering Science Technology (IIEST), Shibpur	341	331	323	97.58	0.77
49	Shanmugha Arts Science Technology Research Academy (SASTRA), Tamil Nadu	332	314	290	92.36	0.69
50	Indian Institute of Technology, Indore	331	330	326	98.79	0.78
51	Jawaharlal Nehru Center for Advanced Scientific Research (JNCASR)	326	321	315	98.13	0.75
52	Indian Institute of Science Education Research (IISER), Kolkata	320	318	307	96.54	0.73
53	Bharathidasan University, Tamil Nadu	317	300	291	97.00	0.69
54	University of Madras, Chennai	317	301	283	94.02	0.68
55	Osmania University, Telangana	313	289	274	94.81	0.65
56	National Institute of Science Education Research (NISER)	307	306	296	96.73	0.71
57	Sri Venkateswara University, Andhra Pradesh	300	277	256	92.42	0.61

58	Jamia Hamdard University, New Delhi	294	275	261	94.91	0.62
59	Tezpur University, Assam	293	288	283	98.26	0.68
60	Pondicherry University, Pondicherry	292	273	258	94.51	0.62
61	Indian Institute of Technology, Hyderabad	286	280	274	97.86	0.65
62	Shivaji University, Maharasthra	285	275	266	96.73	0.63
63	Amrita Vishwa Vidyapeetham University, Tamil Nadu	284	247	219	88.66	0.52
64	King Georges Medical University (KGMU), Lucknow	272	223	171	76.68	0.41
65	Indian Institute of Science Education Research (IISER), Bhopal	261	257	251	97.67	0.60
66	Public Health Foundation of India	257	245	233	95.10	0.56
67	Visvesvaraya National Institute of Technology (VNIT), Nagpur	247	236	226	95.76	0.54
68	Birla Institute of Technology Mesra	240	229	216	94.32	0.52
69	Physical Research Laboratory (PRL), India	239	224	215	95.98	0.51
70	University of Allahabad	232	201	192	95.52	0.46
71	Madurai Kamaraj University, Tamil Nadu	228	218	210	96.33	0.50
72	Homi Bhabha National Institute, Mumbai	228	228	222	97.37	0.53
73	Jawaharlal Institute of Postgraduate Medical Education Research (JIPMER)	225	205	166	80.98	0.40
74	University of Mysore	225	208	194	93.27	0.46
75	University of Kashmir	224	206	191	92.72	0.46
76	Andhra University	222	198	171	86.36	0.41
77	National Institute of Technology (NIT), Durgapur	220	214	209	97.66	0.50
78	Alagappa University, Tamil Nadu	218	210	203	96.67	0.48
79	Maharaja Sayajirao University Baroda	217	212	202	95.28	0.48
80	Sardar Vallabhbhai National Institute of Technology (SVNIT)	217	197	190	96.45	0.45
81	Cochin University Science Technology (CUST), Kerala	217	196	191	97.45	0.46
82	Punjab Agricultural University	213	162	132	81.48	0.31
83	Kalyani University, West Bengal	213	201	189	94.03	0.45
84	National Institute of Technology (NIT), Karnataka	213	207	194	93.72	0.46
85	BOSE Institute	212	208	198	95.19	0.47
86	Jaypee University Of Information Technology	212	208	200	96.15	0.48
87	Institute for Plasma Research IPR	209	209	207	99.04	0.49
88	Central Salt Marine Chemical Research Institute India	207	198	191	96.46	0.46
89	Lucknow University	205	190	175	92.11	0.42
90	Raja Ramanna Centre For Advanced Technology	202	200	195	97.50	0.47
91	L V Prasad Eye Institute	201	169	142	84.02	0.34
92	Inter-University Accelerator Centre	201	199	192	96.48	0.46

[1]Percentage is calculated with respect to number of articles in WoS having a DOI, i.e. column 4.

[2]Percentage calculated with respect to total available Sci-hub article countfor 100 Institutions ofIndia (=41,925)

Open Access through Institutional Repositories

Various ways have been implemented to promote open access in India by several organizations, including creation of Institutional Repositories. The DST-DBT's *sciencecentral.in*,[11] CSIR's *csircentral.net*[12] and ICAR's *krishikosh*[13] are some of the most prominent central institutional repositories in the country. Various government organization have now made it compulsory for their researchers and scientists to submit their research output in relevant IRs. Different funding bodies are also increasingly promoting the cause of public access to public-funded research. The grant award recipients are being encouraged to submit their research outcomes in relevant IRs. Given these initiatives, IRs can play a major role in providing open access to research articles. We have, therefore, tried to analyze the number of research articles that are accessible by the way of being deposited in prominent India IRs.

The *science central. In* repository of the DST-DBT is perhaps the largest central IR in India. It provides an aggregate service that collects the full text and metadata of publications from the DST-DBT institutions as well as outcomes of research funded by them. This IR website says "*at present there are 17 Institutional Repositories hosted at the science central, while 42 institutional repositories are regularly harvested on the same"*. We have conducted a study on the number of articles available through the 24 Institutional repositories that are part of this central repository and found there are total 125,595 publications deposited and available. Table 3 shows the detailed data about number of articles available in each of these 24 repositories. These are very small numbers taking

Table 3: Status of DST-DBT Central Repository (*sciencecentral.in*).

S. No.	Name of Institutional Repository	Period of Records	Total Records
1	IR @ Bharat Immunological and Biologicals Corporation Limited	-	0
2	IR @ Biotechnology Industry Research Assistance Council	-	0
3	IR @ Centre of Innovative and Applied Bioprocessing	2016	2
4	IR @ Centre for DNA Fingerprinting and Diagnostics	1999-2019	877
5	IR @ Department of Biotechnology	-	0
6	IR @ Department of Science and Technology	2014-17	19
7	IR @ Indian Academy of Sciences	1920-2017	98205
8	IR @ Institute of Bioresources and Sustainable Development	2013-14	10
9	IR @ Institute of Life Sciences	2014-15	10
10	IR @ National Agri-Food Biotechnology Institute	2010-15	35
11	IR @ National Brain Research Centre	2003-18	437
12	IR @ National Centre for Cell Science	2012-18	577
13	IR @ National Institute of Animal Biotechnology	2012-14	9
14	IR @ National Institute of Biomedical Genomics	2013-14	10
15	IR @ National Institute of Immunology	2014-15	10
16	IR @ National Institute of Plant Genome Research (NIPGR)	2013-17	772
17	IR @ Rajiv Gandhi Centre for Biotechnology	2011-18	599
18	IR @ Raman Research Institute Digital Repository	2005-15	5947
19	IR @ Regional Centre for Biotechnology	2010-18	117
20	IR @ S.N.Bose National Centre for Basic Sciences	2012-18	1684
21	IR @ SreeChitraTirunal Institute for Medical Sciences and Technology, Trivandrum	2012-13	1800
22	IR @ The Indian Association for the Cultivation of Science	2010-18	7941
23	IR @ Translational Health Science and Technology Institute	2013-14	9
24	IR@Indian Institute of Astrophysics	2004-15	6525
25	**Total Records**		1,25,595

into account the research output volume of India. In order to arrive an estimate, the total publication count of India in Web of Science could be used. A total of 80,9404 articles are there indexed in Web of Science for India for the said period. If we take this as the complete publication volume of India, the number of publications available in this central repository would be only about 15%. However, since it is quite obvious that all research output from India is not indexed in Web of Science and in fact the actual publication volume would be much higher than that, it can be concluded that articles accessible through this IR will be lesser than even 15% of the total volume.

The *csircentral.net* IR is the other big central repository. All the IRs of different labs that are part of CSIR system are regularly harvested by this central harvester. We observe that it has about 30 IR's with a total of about 100,609 articles deposited. Table 4 shows the details of number of articles in each of these 30 IRs. In this case also, taking the Web of Science publication count as volume benchmark for India for the concerned period, it is seen that there are 747,759 papers in Web of Science, out of which less than 13% are available through this IR. Given that in actual practice the total publication volume for India would be much higher than the Web of Science indexed volume, this availability proportion is actually much lesser than 13%.

The third major IR is ICAR system's *krishikosh* repository which stands for Knowledge based Resources Information Systems Hub for Innovations (KRISHI) in agriculture. It consists of Surveys/Experiments/Observational studies, Geospatial data, Learning resources, Publications etc., as it is the centralized data repository of ICAR. The KRISHI community is divided according to subject division matter (SDM). Some of the subject divisions are Agricultural Education (1285), Agricultural Engineering (364), Agricultural Extension (235), Animal Science (1720), Crop Science (4025), Fisheries (2721), Horticulture Science (3677), Natural Resource Management (4252) etc. The numbers in bracket indicate the number of available resources in the repository. Table 5 the details of number of items available in each of the repositories. The total number of records available till date on all the repositories

Table 4: Status of CSIR Institutional Repository (*csircentral.net*).

S.No	Name of Institutional Repository	Period of Records	Total Records
1	IR@AMPRI: CSIR-Advanced Materials and Processes Research Institute, Bhopal	2013-14	764
2	IR@CBRI: CSIR-Central Building Research Institute, Roorkee	2012-16	1078
3	IR@CDRI: CSIR-Central Drug Research Institute, Lucknow	2009-15	1070
4	IR@CECRI: CSIR-Central Electrochemical Research Institute, Karaikudi	2012-19	2639
5	IR@CEERI: CSIR-Central Electronics Engineering Research Institute, Pilani	2013-17	226
6	IR@CFTRI: CSIR-Central Food Technological Research Institute, Mysore	2008-19	9031
7	IR@CGCRI: CSIR-Central Glass and Ceramic Research Institute, Kolkata	2011-18	3692
8	IR@CIMFR: CSIR-Central Institute of Mining and Fuel Research, Dhanbad	2011-19	1894

9	IR@CLRI: CSIR-Central Leather Research Institute, Chennai	2013-16	10
10	IR@CMERI: CSIR- Central Mechanical Engineering Research Institute (CMERI), Durgapur	2016-19	501
11	IR@C-MMACS: CSIR-Centre for Mathematical Modelling and Computer Simulation, Bangalore	2011-14	209
12	IR@CRRI: CSIR-Central Road Research Institute, New Delhi	2017	352
13	IR@CSMCRI: CSIR-Central Salt and Marine Chemicals Research Institute, Bhavnagar,	2011-12	1252
14	IR@IGIB: CSIR-Institute of Genomics and Integrative Biology, New Delhi	2011-17	97
15	IR@IHBT: CSIR-Institute of Himalayan Bioresource Technology, Palampur	2011-18	1038
16	IR@IICB: CSIR-Indian Institute of Chemical Biology, Kolkata	2011-18	2053
17	IR@IICT: CSIR-Indian Institute of Chemical Technology, Hyderabad	2009-14	9337
18	IR@IIIM: CSIR-Indian Institute of Integrative Medicine, Jammu	2011-12	200
19	IR@IIP: CSIR-Indian Institute of Petroleum, Dehradun	2009-18	466
20	IR@IMMT: CSIR-Institute of Minerals and Materials Technology, Bhubaneswar	2009	19
21	IR@IMTECH: CSIR-Institute of Microbial Technology, Chandigarh	2011-19	2199
22	IR@NAL: CSIR-National Aerospace Laboratories, Bangalore	2011-17	6702
23	IR@NCL: CSIR-National Chemical Laboratory, Pune	2018-19	2302
24	IR@NEERI: National Environment Engineering Research Institute	2011-16	321
25	IR@NEIST: CSIR-North East Institute of Science and Technology, Jorhat	2011-12	340
26	IR@NIO: CSIR-National Institute Of Oceanography, Goa	2006-15	4688
27	IR@NISCAIR: CSIR-NISCAIR, New Delhi - ONLINE PERIODICALS REPOSITORY (NOPR)	2011-19	37865
28	IR@NML: CSIR-National Metallurgical Laboratory, Jamshedpur	2011-19	6744
29	IR@NPL: CSIR-National Physical Laboratory, New Delhi	2012-19	2889
30	IR@SERC: CSIR-Structural Engineering Research Centre, Chennai	2016-17	631
	Total Records		1,00,609

Table 5: Status of ICAR Central Repository (*krishikosh*).

S. No.	Name of Institutional Repository	Total Records
1	ICAR-Agricultural Technology Application Research Institutes of all zones together locatedat Bengaluru, Hyderabad, Jabalpur, Jodhpur, Kanpur, Kolkata, Ludhiana, Patna, Umiam, Guwahati and Pune.	235
2	ICAR-Central Agroforestry Research Institute	62
3	ICAR-Central Arid Zone Research Institute	1947
4	ICAR-Central Avian Research Institute	205
5	ICAR-Central Citrus Research Institute	87
6	ICAR-Central Coastal Agricultural Research Institute	163
7	ICAR-Central Inland Fisheries Research Institute	390
8	ICAR-Central Institute for Arid Horticulture	489
9	ICAR-Central Institute for Cotton Research	113
10	ICAR-Central Institute for Post-Harvest Engineering and Technology B4	87
11	ICAR-Central Institute for Research on Buffaloes	42
12	ICAR-Central Institute for Research on Cattle	30
13	ICAR-Central Institute for Research on Cotton Technology	20
14	ICAR-Central Institute for Research on Goats	39
15	ICAR-Central Institute for Women in Agriculture	366
16	ICAR-Central Institute of Agricultural Engineering	148
17	ICAR-Central Institute of Brackish water Aquaculture	841
18	ICAR-Central Institute of Fisheries Education	35
19	ICAR-Central Institute of Fisheries Technology	944
20	ICAR-Central Institute of Freshwater Aquaculture	48
21	ICAR-Central Institute of Sub-tropical Horticulture	154
22	ICAR-Central Institute of Temperate Horticulture	56
23	ICAR-Central Island Agricultural Research Institute	87
24	ICAR-Central Marine Fisheries Research Institute	01
25	ICAR-Central Plantation Crops Research Institute	339
26	ICAR-Central Potato Research Institute	970
27	ICAR-Central Research Institute for Jute and Allied Fibres	126
28	ICAR-Central Research Institute of Dryland Agriculture	594
29	ICAR-Central Sheep and Wool Research Institute	11
30	ICAR-Central Soil Salinity Research Institute	493
31	ICAR-Central Tobacco Research Institute	259
32	ICAR-Central Tuber Crops Research Institute	37
33	ICAR-Directorate of Cashew Research	19
34	ICAR-Directorate of Cold Water Fisheries Research	15
35	ICAR-Directorate of Floricultural Research	38
36	ICAR-Directorate of Groundnut Research	424

37	ICAR-Directorate of Medicinal and Aromatic Plants Research	07
38	ICAR-Directorate of Poultry Research	130
39	ICAR-Directorate of Rapeseed and Mustard Research	34
40	ICAR-Directorate of Weed Research	76
41	ICAR-Indian Agricultural Research Institute	07
42	ICAR-Indian Agricultural Statistics Research Institute	663
43	ICAR-Indian Grassland and Fodder Research Institute	27
44	ICAR-Indian Institute of Agricultural Biotechnology	15
45	ICAR-Indian Institute of Farming Systems Research	101
46	ICAR-Indian Institute of Horticultural Research	526
47	ICAR-Indian Institute of Maize Research	169
48	ICAR-Indian Institute of Millets Research	344
49	ICAR-Indian Institute of Natural Resins and Gums	104
50	ICAR-Indian Institute of Oil Palm Research	74
51	ICAR-Indian Institute of Oilseeds Research	778
52	ICAR-Indian Institute of Pulses Research	37
53	ICAR-Indian Institute of Rice Research	168
54	ICAR-Indian Institute of Seed Science	56
55	ICAR-Indian Institute of Soil and Water Conservation	48
56	ICAR-Indian Institute of Soil Science	56
57	ICAR-Indian Institute of Soybean Research	185
58	ICAR-Indian Institute of Spices Research	93
59	ICAR-Indian Institute of Sugarcane Research	34
60	ICAR-Indian Institute of Vegetable Research	153
61	ICAR-Indian Institute of Water Management	347
62	ICAR-Indian Institute of Wheat and Barley Research	08
63	ICAR-Indian Veterinary Research Institute	08
64	ICAR-National Academy of Agricultural Research and Management	68
65	ICAR-National Bureau of Agricultural Insect Resources	93
66	ICAR-National Bureau of Agriculturally Important Micro-organisms	11
67	ICAR-National Bureau of Animal Genetic Resources	94
68	ICAR-National Bureau of Fish Genetic Resources	447
69	ICAR-National Bureau of Plant Genetics Resources	56
70	ICAR-National Bureau of Soil Survey and Land Use Planning	37
71	ICAR-National Centre for Integrated Pest Management	159
72	ICAR-National Dairy Research Institute	07
73	ICAR-National Institute for Plant Biotechnology	64
74	ICAR-National Institute of Abiotic Stress Management	54
75	ICAR-National Institute of Agricultural Economics and Policy Research	188
76	ICAR-National Institute of Animal Nutrition and Physiology	239
77	ICAR-National Institute of Biotic Stress Management	28
78	ICAR-National Institute of High Security Animal Diseases	107
79	ICAR-National Institute of Natural Fibre Engineering and Technology (earlier NIRJAFT)	05
80	ICAR-National Institute of Veterinary Epidemiology and Disease Informatics	280
81	ICAR-National Organic Farming Research Institute	0
82	ICAR-National Research Centre for Banana	253
83	ICAR-National Research Centre for Grapes	89
84	ICAR-National Research Centre for Litchi	81
85	ICAR-National Research Centre for Pomegranate	05
86	ICAR-National Research Centre on Camel	53
87	ICAR-National Research Centre on Equines	85
88	ICAR-National Research Centre on Integrated Farming	0
89	ICAR-National Research Centre on Meat	237
90	ICAR-National Research Centre on Mithun	112
91	ICAR-National Research Centre on Orchids	13
92	ICAR-National Research Centre on Pig	18
93	ICAR-National Research Centre on Seed Spices	45
94	ICAR-National Research Centre on Yak	06
95	ICAR-National Rice Research Institute	661
96	ICAR-Project Directorate on Foot and Mouth Disease	17
97	ICAR-Research Complex for Eastern Region	31
98	ICAR-Research Complex for NEH Region	243
99	ICAR-Sugarcane Breeding Institute	07
100	ICAR-Vivekananda Parvatiya Krishi AnusandhanSansthan	162
101	Others-Others-AICRP on Pearl Millet Publication	07
102	Others-Others-Data Inventory	00
103	Others-Others-Other	00
104	Others-Others-Publication	262
	Total Records	18,486

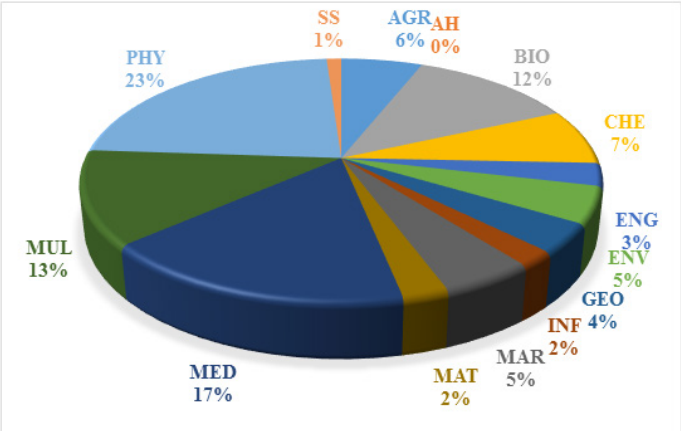

Figure 1: Discipline wise distribution of OA articles for collective data of institutions.

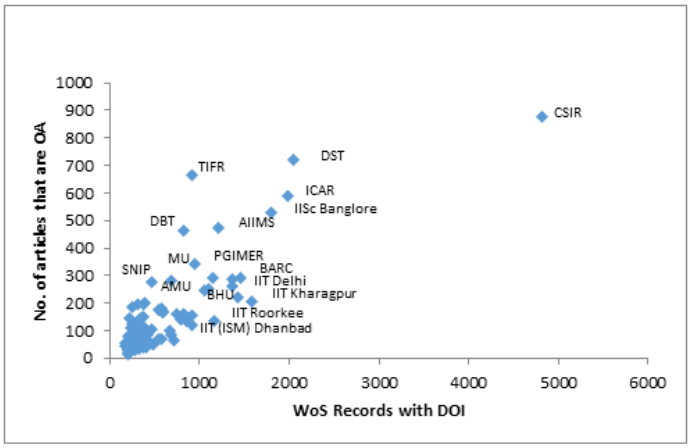

Figure 2: Scatter plot of OA articles vs. total no. of articles for 100 most productive Indian Institutions.

taken together are 18,486, which would again be a very small proportion of Indian research output in the area.

The most recent initiative in the repository culture of open access in India is the creation of IndiaRxiv hosted at indiarxiv.org. However, there are only 67 preprints available as on 16th Feb. 2020, almost 10 months after the launch of the repository in April 2019. This repository is thus yet to gain popularity and momentum of deposits.

CONCLUSION

The paper tried to analyze the research output of 100 most productive institutions (including institution systems) in India for the year 2016 and measured the proportion research output available in open access. The study not only provides important observations on the status and pattern of open access in Indian institutions, but also draw attention to some inhibiting factors/constrains that prevents Indian research outputs from being available in open access mode.

Firstly, it is observed that only 23% of the combined output of the 100 institutions in 2016, as indexed in Web of Science, are available in open access. Gold Open access is the most prevalent type followed by Green and Brown open access. The lack of visibility of articles in hybrid open access may have deeper causal reasons. It has now become common for journals to have a hybrid open access model which acts as a revenue source for them. However, the cost of keeping a paper open access in most of the journals that follow this model is exorbitantly high. Institutions in developed economies have articulated policy that allows researchers to exploit this opportunity to keep their papers open access. The window available for Indian researchers to publish in open access mode is much less than those countries where institutional support is given. Thus, the findings of this study have to be seen in this context. Studies clearly point out and unsurprisingly so that a paper that is in open access has more chance for dissemination across the research community and attract higher citations and recall. Thus, a researcher if provided support would exploit it to publish in journals that follow hybrid open access policy.

Secondly, the open access proportions in the 100 most productive institutions are found to vary significantly, ranging from as low as 7% to as high as 75% of the total published papers. TIFR, Saha Institute of Nuclear Physics, NISER, PRL and PHFI are some of the institutions with high percentage of their articles available in open access. The Engineering and Technology institutions, including several IITs are found to have low percentage of their articles available as open access. A more detailed study may highlight the inhibiting factors behind the lower percentage of articles in open access from engineering and technology institutions.

Thirdly, the paper also found disciplinary variations in open access levels, with disciplines like PHY and MED having a higher proportion of articles available as open access. Papers from ENG and INF disciplines are found to have low open access proportion. Institutions in general having higher open access are primarily those that have Physics or Medicine as their major areas of activity. Interestingly, the geographical location of an institution is not found connected to the level of open access availability of research output of the institutions.

Fourthly, the number of papers available in Sci-Hub from all the 100 institutions (including institution systems) is many times higher than those available in legal open access forms. More than 2/3rd of the institutions in the set have higher than 90% of their papers available for free download in Sci-Hub. These results are in a sense an indication of failure of the legal OA models. Results indicate that there are either deterrents to use of legal OA models or possibly some kind of apathy in researchers to use legal OA models (such as depositing papers in institutional or disciplinary repositories).

Lastly, analysis of major Indian IRs show that the proportion of papers deposited in the IRs is very low and IRs as such could not emerge as a significant source of open access availability of research articles. The analysis of the three major repositories: *Sciencentral*, *csircentral* and *krishikosh* suggest that more efforts are needed to promote the IR culture in India. At the same time a lot needs to be done to understand why IRs are not an attractive medium for Indian researchers to disseminate their research in open access. These efforts could not only be limited to funding agencies but all the research institutions should make it mandatory for their researchers to submit their papers (or pre-print or post-print) to concerned IRs. Individual scientists should also be encouraged and incentivized for submitting their papers to IRs. This would not only help other Indian researchers who do not have access to costly journal subscriptions, but also the researchers themselves as their paper would find more citations and use.

It is quite clear from the study that much more needs to be done to promote open access culture in India, including the provision of incentives to researchers and institutions that help promote open availability of research output from Indian institutions. The study opens up new research questions, possibilities for expanding the scope of research and some important lessons.

One important aspect that this paper could not analyze is the amount of research output from India that is deposited in disciplinary repositories like arXiv or in academic social networks (such as ResearchGate). Academic social networks are now increasingly being used by researchers and also indexed by Web search engines. There are some difficulties in accessing these platforms for an automated crawling which prevented us from analyzing data from them. However, we are working on finding solutions to this and hope to look at this aspect as well in future work.

ACKNOWLEDGEMENT

The authors acknowledge the enabling support provided by the DST-NSTMIS funded project 'Design of a Computational Framework for Discipline-wise and Thematic Mapping of Research Performance of Indian Higher Education Institutions (HEIs)', bearing Grant No. DST/NSTMIS/05/04/2019-20 for this work.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. Urs SR, Raghavan KS. Vidyanidhi: Indian digital library of electronic theses. Communications of the ACM. 2001;44(5):88-9.
2. Kuri R. Information and Knowledge Sharing Through Institutional Repository: An Overview of EPRINTS@IISC International Journal of Digital Library Services. 2014;4(4). ISSN:2250-1142 (Online), ISSN 2349-302X
3. National Knowledge Commission. National Knowledge Commission. Report to the Nation 2006-2009. 2009. Available at: http://14.139.60.153/bitstream/123456789/112/1/National%20Knowledge%20Commission-Report%202006-2009%20.pdf accessed on 20th Aug. 2019.
4. Das AK. Open Access to Knowledge and Information: Scholarly Literature and Digital Librarylnitiatives- The South Asian Scenario. UNESCO Report, New Delhi. 2008.
5. Arunachalam S, Muthu M. Open Access to Scholarly Literature in India- A Status Report (with Emphasis on Scientific Literature). Centre for Internet and Society, Bangalore, India. 2011. available at: https://cis-india.org/openness/publications/open-access-scholarly-literature.pdf
6. DBT and DST Open Access Policy. Policy on Open Access to DBT and DST Funded Research, Department of Biotechnology and Department of Science and Technology, Government of India. 2014. Available at: http://www.dst.gov.in/sites/default/files/APPROVED%20OPEN%20ACCESS%20POLICY-DBT%26DST%2812.12.2014%29_1.pdf accessed on 20-08-2019.
7. Hajjem C, Harnad S, Gingras Y. Ten-year Cross-Disciplinary of the growth of Open Access and how it increases research citation impact. Bulletin of the IEEE Computer Society Technical Committee on Data Engineering. 2005.
8. Björk B. Hybrid open access-A longitudinal study. Journal of Informetrics. 2016;10(4): 919. https://doi.org/10.1016/j.joi.2016.08.002
9. Björk BC. The hybrid model for open access publication of scholarly articles: A failed experiment? Journal of the American Society for Information Science and Technology. 2012;63(8):1496-504.
10. Björk BC. Gold, green and black open access. Learned Publishing. 2017;30:173-5. doi:10.1002/leap.1096
11. Björk BC, Laakso M, Welling P, Paetau P. Anatomy of green open access. Journal of the Association for Information Science and Technology. 2014;65(2):237-50.
12. Björk BC, Welling P, Laakso M, Majlender P, Hedlund T, Guðnason G. Open Access to the Scientific Journal Literature: Situation 2009. PLoS ONE. 2010;5(6):e11273. https://doi.org/10.1371/journal.pone.0011273
13. Björk BC. The open access movement at a crossroad: Are the big publishers and academic social media taking over?. Learned Publishing. 2016;29(2);131-4. https://doi.org/10.1002/leap.1021
14. Archambault É, Amyot D, Deschamps P, Nicol A, Provencher F, Rebout L, *et al*. Proportion of open access papers published in peer-reviewed journals at the European and world levels—1996–2013. European Commission. 2014. Retrieved from http://science-metrix.com/sites/default/files/science-etrix/publications/d_1.8_sm_ec_dg-rtd_proportion_oa_1996-2013_v11.pdf on 25 June 2019.
15. Archambault E, Amyot D, Deschamps P, Nicol A, Rebout L, Roberge G. Proportion of open access peer-reviewed papers at the European and world levels-2004-2011. European Commission. 2013. Retrieved from http://science-metrix.com/pdf/SM_EC_OA_Availability_2004-2011.pdf
16. Piwowar H, Priem J, Larivière V, Alperin JP, Matthias L, Norlander B, *et al*. The state of OA: A large-scale analysis of the prevalence and impact of Open Access articles. Peer J. 2018;6:e4375. https://doi.org/10.7287/peerj.preprints.3119v1 accessed on 20-08-2019
17. Bosman J, Kramer B. Open access levels: A quantitative exploration using Web of Science and oaDOI data (No. e3520v1). PeerJ Preprints. 2018.
18. Kumar V, Mahesh G. Open access repositories in India: A lost opportunity. Current Science. 2017;112(2):210-12.
19. Mallapaty S. Indian scientists launch preprint repository to boost research quality. Nature. 2019. Retrieved from: https://www.nature.com/articles/d41586-019-01082-0
20. Piryani R, Dua J, Singh. Open Access Levels and Patterns in Scholarly Articles from India. Current Science. 2019;117(9):1435-40.
21. Singh VK, Piryani R, Srichandan SS. The Case of Significant Variations in Gold-Green and Black Open Access: Evidence from Indian Research Output. Scientometrics, in Press. 2020. DOI: 10.1007/s11192-020-03472-y
22. Das AK, Dutta B. Open Science in Addressing Responsible Research and Innovation: Evidence from India and other Countries. Journal of Scientometric Research. INDIALICS Special Issue. 2020.
23. Rupika, Uddin, A., Singh, V.K. Measuring the university-industry-government collaboration in Indian Research Output. Current Science. 2016;110(10):1904-9.
24. Greshake-Tzovaras B. "Looking into Pandora's box: The content of Sci-Hub and its usage". F1000Research. 2017;6:541. https://doi.org/10.12688/f1000research.11366