

Assessing Open Access Friendliness of National Institutes of Technology (NITs): A Data Carpentry Approach

Abhijit Roy* and Parthasarathi Mukhopadhyay

Department of Library and Information Science, University of Kalyani, Kalyani, Nadia - 741 235, India

**E-mail: royabhijit75@gmail.com*

ABSTRACT

This research study aims to measure the Open Access (OA) friendliness of National Institutes of Technology (NITs) of India that are listed in the overall category of NIRF (National Institutional Ranking Framework), 2021 by taking into consideration four important OA parameters – i) OA publication share; ii) OA licensing scenario; iii) citation impact of OA publications; and iv) altmetric scores of OA publications. It deals with 64,485 publications of the selected 11 NITs during the period from 2012 to 2021 (10 years), citations received by these publications (5,42,638 citations), and altmetric attention scores of the documents (5,213 publications) during the period under study. A data carpentry tool, namely OpenRefine, and open access bibliographic/citation data sources such as Unpaywall, Dimensions, and Altmetric.com have been deployed to accomplish this large-scale study for ranking NITs by their Open Access Friendliness (OAF). The OAF indicator, as applied in this study, is a distributed weightage based 100-point scale built on top of the aforesaid OA parameters. The ranking framework shows that Sardar Vallabhbhai National Institute of Technology, Surat (est. in 1961) has achieved the top position with a score of 52.12 (out of 100), but in totality only 3 NITs (out of the selected 11 NITs) crossed the 50 per cent mark in the adapted OAF scale.

Keywords: Open access; Open access indicators; Open access friendliness; Data carpentry; OpenRefine

1. INTRODUCTION

The concept of Open Access (OA) has taken the scholarly world by storm, and India is no exception to this emerging trend of OA scholarly communication¹. Most of the higher educational institutes in India have adopted OA policies, and subsequently received the advantages in research impact. The public data dump of DOAJ (Directory of Open Access Journals) shows that a total of 325 OA journals (Gold OA) are presently published from India (as on August 31, 2022) and India occupies the 18th position out of the 130 represented countries in DOAJ. Similarly, other global databases like OpenDOAR (<https://v2.sherpa.ac.uk/opensoar>) and ROAR (<http://roar.eprints.org/>) have reported an increasing number of OA repositories (Green OA) from India (105 in OpenDOAR and 135 in ROAR, respectively as on August 31, 2022). But till date, there are only a few initiatives to measure OA support in India at the institutional level. This research study is a part of measuring the OA friendliness of a given set of institutes (here, NITs in India) on the basis of a framework proposed by an earlier research work of the coauthor of this paper.

2. RELATED WORKS AND PURPOSE OF THE STUDY

Open access initiatives and the development of the open knowledge movement in India have been analysed by a few

researchers to different degrees¹⁻⁹. A group of researchers have attempted to measure open access publications at the country/regional level^{1,6,7,10,11}. Other researchers like Alperin, Babini, Gomez and Maddi¹²⁻¹⁵ have attempted to quantify open access indicators based on availability institutional repository, open access colours, adaptation of creative commons license and so on. A team of researchers under the leadership of Robinson-Garcia^{16,17} has conducted a large-scale global study of 963 universities (including India) regarding open access uptake. They have developed a framework to measure the impact of OA publications of the selected 963 universities for five years (2014–2017) on the basis of Web of Science publications data and Unpaywall acted as a data source for OA status. A new scientometric method to handle large-scale publications and citations data for measuring OA friendliness has recently been proposed by Mukhopadhyay¹⁸ using the data carpentry method. This research study ranked NIRF listed Indian Institutes of Technology (IITs) by their OA friendliness based on OA publications, OA licences, citations received by OA publications and altmetric scores, and found that the newly established IITs are more OA friendly than the older and established IITs.

The purpose of this research is to assess the OA friendliness of National Institutes of Technology (NITs) enlisted in the overall category of the National Institutional Ranking Framework (NIRF) 2021 (Table 1) using a 100-point weightage-based scale, taking into account four major areas of OA as proposed by Mukhopadhyay¹⁸: (i) OA publication

Table 1. NITs in NIRF (2017-2021) – Arranged by ranks in NIRF, 2021(Overall category)

| Name of NITs | Founded | Est. as NIT | Rank in NIRF (2017-2021) | | | | |
|---|---------|-------------|--------------------------|------|------|------|------|
| | | | 2017 | 2018 | 2019 | 2020 | 2021 |
| National Institute of Technology Tiruchirappalli | 1964 | 2002 | 34 | 31 | 24 | 24 | 23 |
| National Institute of Technology Karnataka | 1960 | 2002 | 65 | 57 | 53 | 33 | 32 |
| National Institute of Technology Rourkela | 1961 | 2002 | 46 | 42 | 38 | 32 | 41 |
| Visvesvaraya National Institute of Technology | 1960 | 2002 | NL | NL | NL | NL | 54 |
| National Institute of Technology Warangal | 1959 | 2002 | 82 | 78 | 61 | 46 | 59 |
| National Institute of Technology Durgapur | 1960 | 2002 | NL | NL | 93 | 96 | 71 |
| Malaviya National Institute of Technology | 1963 | 2002 | NL | NL | NL | 71 | 72 |
| Dr. B. R. Ambedkar National Institute of Technology | 1987 | 2002 | NL | NL | NL | NL | 78 |
| Motilal Nehru National Institute of Technology | 1961 | 2002 | NL | NL | NL | 93 | 88 |
| National Institute of Technology Silchar | 1967 | 2002 | NL | NL | NL | 94 | 93 |
| Sardar Vallabhbhai National Institute of Technology | 1961 | 2002 | NL | NL | NL | NL | 98 |

(NL: Not listed in NIRF in that year)

share; (ii) OA licensing scenario; (iii) citations received by OA publications; and (iv) altmetric scores obtained by the published papers. These major areas are again subdivided into different divisions with respective weightages to design a framework for assessing the OA friendliness of NITs.

3. METHODOLOGY

The methodology of this research study is extensively based on an open-source data wrangling tool—OpenRefine. It collects required datasets from the domain-specific sources available under the Open Data Commons Open Database License (ODbL). The ranking framework centres around four main factors, namely: OA publications share, OA license scenario, OA citations share, and OA altmetric scores share. The deployed framework includes a total of nine factors under these selected areas. The distributions of weightage have been set on the basis of SWOC analysis of the open access scenario in India as proposed by Mukhopadhyay¹⁸. The framework was tested for ranked NITs in India (11 in total - see Table 1) using publication data (64,485 publications by 11 NITs), citation data (5,42,638 citations received by these papers), and altmetric data (for 5,213 publications). The overall methodology may broadly be divided into two groups – a) tuning the OA friendliness ranking framework; and b) gathering, extracting

and calculating data sets as required for the ranking framework through data carpentry.

3.1 Ranking Framework

As stated in section 2, this research study used Mukhopadhyay's¹⁸ proposed ranking framework, which demonstrated test results for a massive number of publications and citation datasets from Indian Institutes of Technology (IITs) (16 in total). The ranking framework of Mukhopadhyay is modified slightly in terms of weightages and factors to suit the nature of citation data sets of NITs under consideration and includes four factors of OA with nine parameters under these four factors (Table 2).

The scopes of these nine parameters under the four major OA factors are as follows.

3.1.1 OA Share

It represents for a given NIT - total OA publications during the period divided by the total publications (with OA status from Unpaywall) during the period, and the score is multiplied by the weightage of 25.

3.1.2 Green and Gold OA Share

Green and Gold OA are considered important for this

Table 2. Components of the ranking framework

| Areas (weightage) | Groups within the areas (with distributed weightage) | | |
|--|--|--|---|
| OA Publications (Area I - weightage 50%) | OA share (Group weightage: 25%) | Gold & Green share (Group weightage: 15%) | Repository share (Group weightage:10%) |
| OA Licensing (Area II - weightage 30%) | OA license share (Group weightage: 20%) | Gold & Green license share (Group weightage: 10%) | |
| OA impact: Citations (Area III - weightage 10%) | OA citation share type I (Group weightage: 5%) | OA citation share type II (Group weightage: 5%) | |
| OA impact: Altmetric (Area IV - weightage 10%) | OA altmetric share type I (Group weightage: 5%) | OA altmetric share type II (Group weightage: 5%) | |

study because these two routes are more legalistic than the other routes (Bronze & Hybrid). It measures, for a given NIT - total gold & green OA publications during the period divided by the total OA publications during the period and multiplied by the weightage of 15.

3.1.3 Repository Share

It is calculated by summing up the total green OA papers of an institute available through repositories (as best OA location from Unpaywall) and dividing the sum by the total green OA papers made available by that institute, and then multiplied by the weightage value of 10.

3.1.4 OA License Share

It indicates - total OA publications with formal OA licenses (during the period) divided by the total OA publications during the period and then multiplied by the weightage value of 20.

3.1.5 Gold and Green License Share

In view of the increasing importance of OA licensing for green and gold OA, this ranking framework has given an additional weightage to measure: total gold & green OA publications with formal OA licenses (during the period) divided by the total gold & green publications during the period, and then multiplied by the allocated weightage value of 10.

3.1.6 OA Citation Share Type I

It is a ratio of citations received by all publications of a given NIT (A) and citations received by only OA papers (B)

during the period of study, and then multiplied by the allocated weightage value for the group, i.e., $(B \div A * 5)$.

3.1.7 OA Citation Share Type II

The average citation value is utilised here as a cutoff point. It first sums up all citations received by only those publications that have citations greater than ($>$) the average citation value (A), and then sums up all citations received by OA publications with citations greater than the average citation value (B). The group value is obtained by the formula: $(B \div A * 5)$

3.1.8 OA Altmetric Share Type I

This factor is important in view of the increasing influence of socio-academic web spaces on the scholarly communication process. It sums up the altmetric attention scores received by the OA publications of an institution (B), and is divided by the sum of altmetric attention scores of all publications of the institute (A). Both 'A' and 'B' are calculated during the period of study, and the ratio is finally multiplied by the allocated weightage value for the group - $(B \div A * 5)$

3.1.9 OA Altmetric Share Type II

It measures the share of OA in publications with high-value altmetric scores. This parameter is calculated by summing up the altmetric attention scores received by all those publications (OA+non-OA) that have scores greater than ($>$) the average altmetric attention score (C), and then summing up the altmetric attention scores of OA publications only with the altmetric score greater than the average altmetric attention score (D). The group value is obtained by the formula - $(D \div C * 5)$.

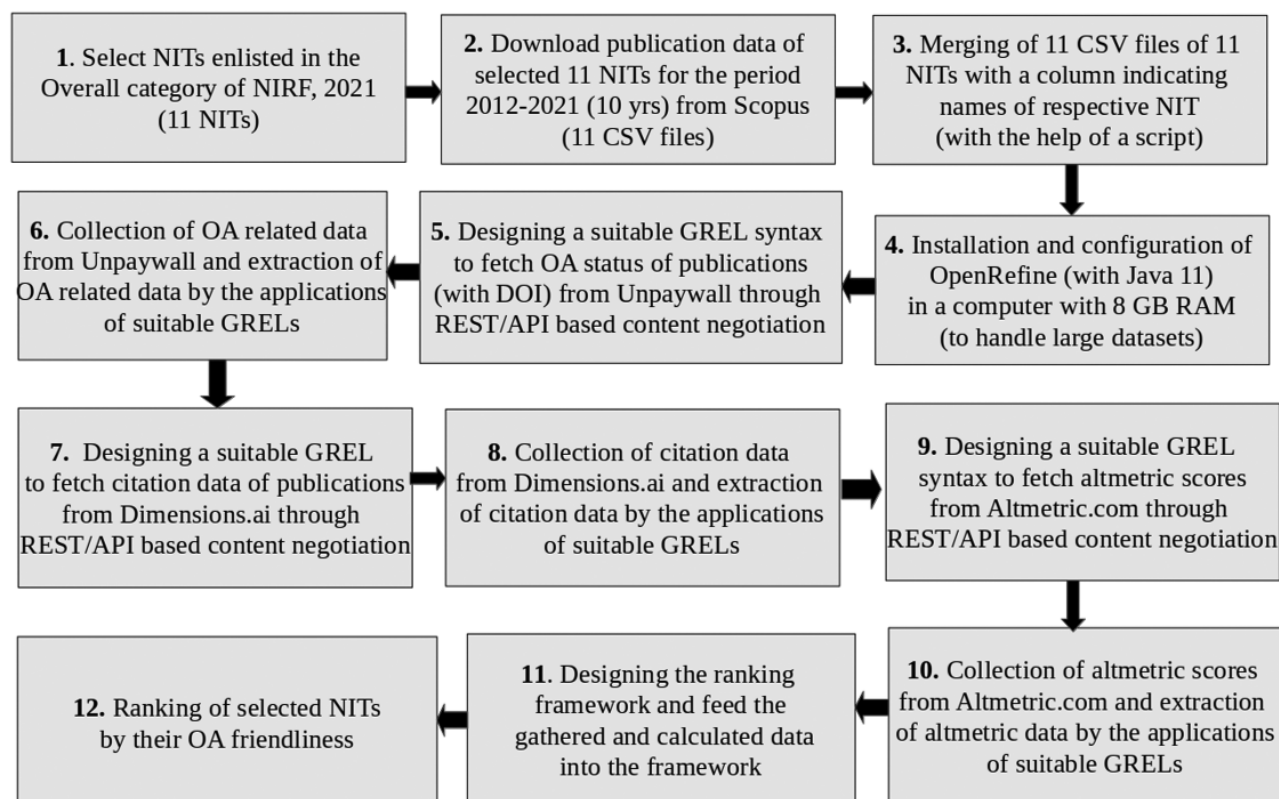


Figure 1. Illustration of steps for data carpentry.

Table 3. Primary dataset (arranged by number of publications in descending order)

| Name of NIT | Total publication (2012-2021) | Publication with DOI | Rank in NIRF 2021 |
|---|-------------------------------|----------------------|-------------------|
| National Institute of Technology Rourkela | 10,392 | 9,974 | 41 |
| National Institute of Technology Tiruchirappalli | 8,369 | 7,920 | 23 |
| National Institute of Technology Karnataka | 7,392 | 6,832 | 32 |
| National Institute of Technology Warangal | 5,571 | 5,226 | 59 |
| National Institute of Technology Durgapur | 5,476 | 5,163 | 71 |
| Malaviya National Institute of Technology | 5,394 | 5,151 | 72 |
| Motilal Nehru National Institute of Technology | 4,801 | 4,517 | 88 |
| Sardar Vallabhbhai National Institute of Technology | 4,892 | 4,485 | 98 |
| National Institute of Technology Silchar | 4,427 | 4,225 | 93 |
| Visvesvaraya National Institute of Technology | 4,280 | 4,033 | 54 |
| Dr. B. R. Ambedkar National Institute of Technology | 3,491 | 3,232 | 78 |
| Total of publications | 64,485 | 60,758 | |

Table 4. REST/API based content negotiation for ODbL data sources

| Content negotiation (REST/API syntax) | No. of queries sent | Responses received |
|--|---------------------------------|--|
| "https://api.unpaywall.org/v2/" + value + "?email=<your-mail-id-here>" | | 60,092 98.90 % of publications with DOI |
| "https://metrics-api.dimensions.ai/doi/" + value | 60,758 publications with DOI | 59,920 98.62 % of publications with DOI |
| "https://api.altmetric.com/v1/doi/" + value + "?key=Your-Key-Here" | | 5,213 8.58 % of publications with DOI |

Note: value is DOI for all datasets and Altmetric.com requires an API key for large-scale data fetching

Table 5. Applications of GRELS in extracting information from JSON datasets (Unpaywall as an example)

| Response from Unpaywall in JSON | GREL for data extraction | Extracted data |
|---|---|---|
| "genre": "journal-article", "published_date": "2011-12-01", "year": 2011, | value.parseJson().is_oa | True |
| "journal_name": "Plant Methods", | value.parseJson().journal_is_in_doaj | True |
| "journal_issns": "1746-4811", | value.parseJson().oa_locations[0].license | cc-by |
| "journal_issn_l": "1746-4811", | value.parseJson().journal_is_oa | True |
| "journal_is_oa": true, | value.parseJson().has_repository_copy | True |
| "journal_is_in_doaj": true, | value.parseJson().publisher | Springer Science and Business Media LLC |
| "publisher": "Springer Science and Business Media LLC", | value.parseJson().oa_status | Gold |
| "is_oa": true, | value.parseJson().journal_is_in_doaj | True |
| "oa_status": "gold", | | |
| "has_repository_copy": true, | | |
| "license": "cc-by", | | |
| "version": "publishedVersion", | | |
| "host_type": "publisher", ... | | |
| | | |

3.2 Data Carpentry Steps

The nine parameters as included in the ranking framework require three groups of raw datasets, namely: a) OA status of a publication; b) citations received by a publication; and c) altmetric score of a publication. The data carpentry method is extremely useful here to gather required datasets automatically from different ODbL-based open access data sources in view of the large number of publications in the primary dataset. The

data carpentry steps have four major prerequisites: a) a paper must have an identifier, preferably DOI; b) selected domain-specific data sources must support content negotiation through REST/API; c) a REST/API call must supply dataset, preferably in JSON (Java Script Object Notation – a light-weight data exchange format); and d) the data wrangling tool in use must support precise extraction of required data from the JSON dataset as obtained. The selected open-source data wrangling

Table 6. Calculation of scores for Area I & Area II for two selected NITs

| Area I: OA publications (weightage 50) & Area II: OA licensing (weightage 30) | | | | | | | |
|--|--|-----------------------|--------------|--|--|-----------------------|--------------|
| Factors | | Values (round) | | Factors | | Values (round) | |
| SL | Element | NITT | SVNIT | SL | Element | NITT | SVNIT |
| A | Total publications | 8,369 | 4,892 | E | Total Gold & Green OA | 800 | 584 |
| B | Total publications with DOI | 7,920 | 4,485 | F | Total Green OA | 170 | 71 |
| C | Publications with close/open status from Unpaywall | 7,847 | 4,432 | G | Green OA available via repositories (as the best OA locations) | 159 | 64 |
| D | Total OA publications | 1,064 | 766 | H | Licensed OA resources | 547 | 447 |
| Factor 1: OA share ($D \div C * 25$) | | 3.39 | 4.32 | K | Licensed Gold & Green OA | 459 | 375 |
| Factor 2: Gold & Green OA share ($E \div D * 15$) | | 11.28 | 11.44 | Factor 4: OA license share ($H \div D * 20$) | | 10.28 | 11.67 |
| Factor 3: Repository share ($G \div F * 10$) | | 9.35 | 9.01 | Factor 5: Gold & Green license share ($K \div E * 10$) | | 5.74 | 6.42 |
| Area I scores | | 24.02 | 24.77 | Area II scores | | 16.02 | 18.09 |

Table 7. Calculation of scores for Area III & Area IV for two selected NITs

| Area III: OA impact - citations (weightage 10) & Area IV: OA impact - altmetric (weightage 10) | | | | | | | |
|---|--|---------------|--------------|---|--|---------------|--------------|
| Factors | | Values | | Factors | | Values | |
| SL | Element | NITT | SVNIT | SL | Element | NITT | SVNIT |
| A1 | Total publication with citation status | 7,844 | 4,432 | K | Total publications with altmetric score | 654 | 346 |
| A2 | Total publications with citation ≥ 1 | 6,026 | 3,307 | L | Sum of altmetric scores for all publications with score > 0 | 1,231 | 282 |
| B | Total citations received by all resources that are having citations status | 78,574 | 47,107 | M | Average altmetric score publication ($AAvg = L/K$) | 1.88 | 0.82 |
| C | Average citations per publication ($Avg = B/A1$) | 10 | 11 | N | Number of OA publications with altmetric score | 157 | 97 |
| D | OA publications with citation ≥ 1 | 768 | 561 | O | Sum of altmetric scores for OA publications with score > 0 | 404 | 185 |
| E | Citations received by OA resources with citation ≥ 1 | 8,712 | 12,275 | P | All publications with altmetric score $> AAvg$ ($AAvg$ is the average altmetric score here) | 159 | 168 |
| F | All publications with citation $> Avg$ (Avg is the average citation value 'C' here) | 1,992 | 910 | Q | Sum of altmetric scores for all publications with score $> AAvg$ | 923 | 280 |
| G | Citations received by all resources with citation $> Avg$ | 63,248 | 37,478 | R | Number of OA publications with altmetric score $> AAvg$ | 42 | 37 |
| H | OA Publications with citation $> Avg$ | 228 | 131 | S | Sum of altmetric scores for OA publications with score $> AAvg$ | 333 | 184 |
| J | Citations received by OA resources with citation $> Avg$ | 6,656 | 10,504 | Factor 8: OA altmetric share type I ($O \div L * 5$) | | 1.64 | 3.28 |
| Factor 6: OA citation share type I ($E \div B * 5$) | | 0.55 | 1.30 | Factor 9: OA altmetric share type II ($S \div Q * 5$) | | 1.80 | 3.27 |
| Factor 7: OA citation share type II ($J \div G * 5$) | | 0.53 | 1.40 | | | | |
| Area III scores | | 1.08 | 2.70 | Area IV scores | | 3.44 | 6.56 |

tool (OpenRefine), and selected ODbL-supported data sources (Unpaywall for OA status; Dimensions.ai for citation status; and Altmetric.com for altmetric score status) are compatible with the four primary predispositions of the process. The total publications of all the 11 NITs included in this study during the period 2012–2021 is 64,485, of which 60,758 publications have DOI (94.22 % to be exact – see Table 3), and therefore the primary dataset is also quite comprehensive to obtain the OA friendliness scenario for NITs. An illustrative view of the entire data carpentry steps is included to support a lucid representation of the processes (Fig. 1).

3.2.1 Primary Dataset

There are 31 NITs in the country but only 11 NITs are included in the NIRF, 2021 (Overall category – rank band I:1–100). The publications of these 11 selected NITs (as obtained from Scopus) were merged into a single CSV file with the help of a suitable script (Table 3) as the primary dataset for this study. It shows that 3,727 (5.78 %) publications from selected NITs as available from Scopus are without DOIs, and therefore cannot be considered for the data carpentry steps as illustrated in Fig. 1.

3.2.2 Data Wrangling

The data wrangling involves steps and methods for obtaining OA status, citation status, and altmetric attention score status from selected ODbL-based data sources.

3.2.2.1 OA Status

The OA status of publications from the primary dataset has been fetched from Unpaywall through a suitable REST/API syntax (Table 4). Unpaywall has provided OA status for 60,092 publications, out of a total of 60,758 publications, with DOI. Only 666 (1.10 %) publications with DOI did not receive a response from Unpaywall and were therefore excluded from this study. In summary, OA status includes 60,092 publications, i.e., 93.18 per cent of the total publications and 98.90 per cent of the publications with DOI from 11 NITs.

3.2.2.2 Citation Status

The citation status datasets were collected from Dimensions.ai by using REST/API based content negotiation (Table 4), and it shows that a total of 59,920 publications out of 60,758 publications (98.62 %) have received citation status from Dimensions.ai.

3.2.2.3 Altmetric Data

The altmetric attention scores for the publication were obtained from Altmetric.com through a REST API call (Table 4). It shows that 8.58 per cent of the total publications by 11 NITs (5,213 out of 60,758 publications with DOI) have received altmetric scores.

3.2.3 Data Extraction

All of these datasets provide responses in JSON format, and the required information from the JSON datasets (as obtained by data wrangling) needs to be extracted for further study. Mukhopadhyay¹⁸ has provided methods for extracting bibliographic and bibliometric data from JSON using GREL (General Refine Expression Language) applications in a series of publications on library carpentry^{19–21}. Table 5 shows the applications of GREL syntaxes to extract required information from JSON datasets.

4. RESULTS

The datasets obtained in JSON format, extracted using appropriate GRELS, and calculated by taking into account calculated datasets for two NITs—NIT Tiruchirappalli (NITT - NIRF 2021 overall rank 23) and the Sardar Vallabhbhai National Institute of Technology (SVNIT - NIRF 2021 overall rank 98) – the highest and the lowest rank holders in NIRF, 2021 Overall category—are explained with two tables (Table 6 for areas I & II, and Table 7 for areas III & IV).

Table 6 shows that NIT Tiruchirappalli is much ahead of Sardar Vallabhbhai National Institute of Technology in terms of absolute numbers for all factors, but Sardar Vallabhbhai National Institute of Technology has done better in terms of

Table 8. Ranked list of 11 top NITs by OAFI

| Name of NITs | Area I (50) | Area II (30) | Area III (10) | Area IV (10) | OAF (100) | Rank |
|---|----------------|-----------------|------------------|-----------------|--------------|------|
| Sardar Vallabhbhai National Institute of Technology | 24.77 | 18.09 | 2.70 | 6.56 | 52.12 | 1 |
| Motilal Nehru National Institute of Technology | 23.96 | 17.89 | 2.12 | 7.41 | 51.37 | 2 |
| National Institute of Technology Warangal | 24.60 | 19.70 | 1.92 | 3.96 | 50.18 | 3 |
| Malaviya National Institute of Technology | 23.33 | 17.04 | 1.86 | 6.14 | 48.37 | 4 |
| National Institute of Technology Rourkela | 24.32 | 15.87 | 2.07 | 5.81 | 48.07 | 5 |
| Visvesvaraya National Institute of Technology | 24.50 | 16.86 | 1.85 | 4.05 | 47.26 | 6 |
| National Institute of Technology Silchar | 23.70 | 16.87 | 1.07 | 4.97 | 46.61 | 7 |
| Dr. B. R. Ambedkar National Institute of Technology | 23.92 | 16.78 | 1.20 | 3.24 | 45.13 | 8 |
| National Institute of Technology Karnataka | 24.10 | 17.55 | 1.45 | 1.98 | 45.08 | 9 |
| National Institute of Technology Tiruchirappalli | 24.02 | 16.02 | 1.08 | 3.44 | 44.56 | 10 |
| National Institute of Technology Durgapur | 23.07 | 16.19 | 1.75 | 3.09 | 44.11 | 11 |

ratio. The same trend is also visible for areas III & IV related to OA citation share and OA altmetric share, respectively (Table 7).

The OA Friendliness Indicator (OAFI) is a sum of Area I (three factors related to OA publication share), Area II (two factors related to licensing share), Area III (two factors related to OA citation share), and Area IV (two factors related to OA altmetric share). It shows Sardar Vallabhbhai National Institute of Technology's NIRF 2021 overall ranking is 98, but it adapted to open access friendliness, and its OAF score is 52.12 (out of 100). On the other hand, the National Institute of Technology, Tiruchirappalli's NIRF 2021 overall ranking is 23, but its OAF score is 44.56 (out of 100) and it is less OA-friendly than the Sardar Vallabhbhai National Institute of Technology. The ranking list of 11 NITs in terms of OAFI shows similar trends (Table 8). If we closely observe the OA culture (sharing, repository archiving, and licensing of publications) of 11 NITs, we find that the Sardar Vallabhbhai National Institute of Technology has acquired 24.77 out of 50 in area I (OA publications share) and 2.70 out of 10 in area III (OA citation share) while the National Institute of Technology, Warangal has obtained the highest value 19.70 (out of 30) in area II (OA license share), and the Motilal Nehru National Institute of Technology has achieved the top position (7.41 out of 10) in area IV (OA altmetric share score).

As evident from the final result set (Table 8), only 3 out of 11 ranked NITs crossed the mark of 50 in the deployed 100-point scale for measuring OA friendliness, and for almost all NITs under study, the share of OA publications and OA citation share are (areas I & III) low, while OA license share and OA altmetric share (areas II & IV) quite encouraging.

5. CONCLUSION

When will everything be open access? A blog post reporting global OA status claimed in 2019 that if we consider the present pace of OA at the global-scale, more than 50 per cent of research publications are currently available in different OA routes and by 2040 almost all publications will be available as OA knowledge objects (<https://blog.impactstory.org/oa-by-when/>). This report is based on a groundbreaking research study that was conducted using the oaDOI dataset (an earlier version of Unpaywall)²². According to the dashboards service of *wizdom.ai*, India is now ranked 10th in the world in terms of the number of OA publications (985K OA publications and 33 per cent OA share as on August 31, 2022)²³. But to date, no tools or services are available to measure OA support at the institutional level. The development of OA culture and OA policies in India across educational and research institutes requires data related to different OA facets like OA share, OA licenses and OA impact. This research study provides a methodology toward this direction based on open-source software and ODbL-based data sources, and thereby, can be applied to measure the OA performance of a given institute easily. The same methodology can be extended to calculate and compare OA support for other groups of institutes (like state universities, central universities, research organisations, etc.) as listed in different editions of NIRF. Moreover, the NIRF

method may take into consideration the OA performance of a given institute in ranking to boost OA culture in the country.

REFERENCES

1. Roy, B.K.; Biswas, S.C. & Mukhopadhyay, P. Open access to scholarly information in India: Trends and developments. *Int. Res. J. Libr. Inf. Sci.*, 2012, **2**(1), 89-101. https://www.academia.edu/11391575/Open_Access_to_scholarly_information_in_India_Trends_and_Developments?from=cover_page (Accessed on 15 February 2022).
2. Budapest open access initiative. Read the declaration. <https://www.budapestopenaccessinitiative.org/read/> (Accessed on 15 February 2022).
3. Chakravarty, R. & Mahajan, P. Open access journals initiatives in India. *Int. J. Inf. Dissem. Technol.*, 2011, **1**, 9-12. <https://www.ijidt.com/index.php/ijidt/article/viewFile/158/158> (Accessed on 15 February 2022).
4. Fernandez, L. Open access initiatives in India - An evaluation. *Partnersh. Can. J. Libr. Inf. Pract. Res.*, 2006, **1**. doi: 10.21083/partnership.v1i1.110.
5. Keisham, S. & Sophiarani, S. Open access journal and open access initiatives in India. 2008. <https://ir.inflibnet.ac.in/bitstream/1944/1134/1/19.pdf> (Accessed on 15 February 2022).
6. Nazim, M. Analysing open access uptake by academic and research institutions in India. *DESIDOC J. Libr. Inf. Technol.*, 2021, **41**, 108–115. doi: 10.14429/djlit.41.02.16324.
7. Piryani, R.; Dua, J. & Singh, V.K. Open access levels and patterns in scholarly articles from India. *Curr. Sci.*, 2019, **117**, 1435–1440. https://web.archive.org/web/20220224025111id_/https://www.currentscience.ac.in/Volumes/117/09/1435.pdf (Accessed on 15 February 2022).
8. Sahu, S.K. & Arya, S.K. Open access practices in India. *Libr. Hi Tech News*, 2013, **30**, 6–12. doi: 10.1108/LHTN-03-2013-0011.
9. Sawant, S. The current scenario of open access journal initiatives in India. *Collect. Build.*, 2009, **28**, 159–163. doi: 10.1108/01604950910999819.
10. 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. *Eur. Comm.*, 2013. https://science-metrix.com/pdf/SM_EC_OA_Availability_2004-2011.pdf (Accessed on 15 February 2022).
11. Archambault, É.; Amyot, D.; Deschamps, P.; Nicol, A.; Provencher, F.; Rebout, L. & Roberge, G. Proportion of open access papers published in peer-reviewed journals at the European and world levels—1996–2013. 2014. <https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1007&context=scholcom> (Accessed on 15 February 2022).
12. Alperin, J.P.; Packer, A.; Aguado-López, E.; Becerril-García, A.; Babini, D.; Archuby, G.; Carrizo, V.; García, D.A.; Higa, S. & Spano, D. Open access indicators and

- scholarly communications in Latin America. CLACSO, 2014.
<http://eprints.rclis.org/25122/2/Open%20Access%20indicators%20LAC%20-%20book%202014.pdf> (Accessed on 15 February 2022).
13. Gomez, N.; Bustos-Gonzalez, A.; Santillan-Aldana, J. & Arias, O. Open access indicators and information society: The Latin American case. *OCLC Syst. Serv. Int. Digit. Libr. Perspect*, 2009, **25**, 82–92.
doi: 10.1108/10650750910961884.
 14. Maddi, A. Construction of a normalised open access indicator (NOAI). 2019. <https://hal.archives-ouvertes.fr/hal-02328158/file/DT-CEPN-2019-08.pdf> (Accessed on 15 February 2022).
 15. Maddi, A. Measuring open access publications: A novel normalised open access indicator. *Scientometrics*, 2020, **124**, 379–398.
doi: 10.1007/s11192-020-03470-0.
 16. Robinson-Garcia, N.; Costas, R. & van Leeuwen, T.N. Indicators of open access for universities. arXiv preprint arXiv:1906.03840, 2019. <https://arxiv.org/ftp/arxiv/papers/1906/1906.03840.pdf> (Accessed on 15 February 2022).
 17. Robinson-Garcia, N.; Costas, R. & Leeuwen, T.N. van. Open access uptake by universities worldwide. *Peer J.*, 2020, **8**, (e9410).
doi: 10.7717/peerj.9410.
 18. Mukhopadhyay, P. How green is my valley? Measuring open access friendliness of Indian Institutes of Technology (IITs) through data carpentry. in *Panorama of open access: Progress, practices & prospects* Ess Ess, 2022, 67–89.
doi: 10.5281/zenodo.6511080.
 19. Mukhopadhyay, P.; Mitra, R. & Mukhopadhyay, M. Library carpentry: Towards a new professional dimension (Part I—Concepts and case studies). *SRELS J. Inf. Manage.*, 2021, **58**(2), 67–80.
doi: 10.17821/srels/2021/v58i2/159969.
 20. Mukhopadhyay, P. & Mukhopadhyay, M. Library Carpentry: Towards a new professional dimension (Part II – Automatic authority control to enhance retrieval). *SRELS J. Inf. Manage.*, 2021, **58**(3), 135–155.
doi:10.17821/srels/2021/v58i3/163890.
 21. Mukhopadhyay, P. & Mitra, R. Library carpentry: Towards a new professional dimension (Part III – Data reconciliation, named entity recognition and advanced utilities). *SRELS J. Inf. Manage.*, 2021, **58**(5), 287–303.
doi: 10.17821/srels/2021/v58i5/166770
 22. Piwowar, H.; Priem, J.; Larivière, V.; Alperin, J.P.; Matthias, L.; Norlander, B.; Farley, A.; West, J. & Haustein, S. The state of OA: A large-scale analysis of the prevalence and impact of open ccess articles. *Peer J.*, 2018, **6**, (e4375).
doi: 10.7717/peerj.4375.
 23. Wizdom.ai. OA output by countries. <https://www.wizdom.ai/dashboards/open-access>. (Accessed on August 31, 2022).

CONTRIBUTORS

Mr Abhijit Roy is an MPhil scholar of the Department of Library Science, University of Kalyani. He has carried out review of literature, overview of initiatives, data entry works, data fetching activities, testing & debugging of the methodology and drafting of this paper.

Mr Parthasarathi Mukhopadhyay is a Professor in Department of Library Science, University of Kalyani. His contributions to the current study are: Developing the OA friendliness ranking framework, designing GREs for data fetching and data extraction in OpenRefine, and formulation of the methodology.

© 2022. This work is licensed under

[http://creativecommons.org/licenses/by-nc-nd/2.5/in\(the “License”\)](http://creativecommons.org/licenses/by-nc-nd/2.5/in(the%20License)).

Notwithstanding the ProQuest Terms and Conditions, you may use this content
in accordance with the terms of the License.