

# **Unmasking Unconventional Impact: Assessing Research Contributions of Private Higher Education Institutions (HEIs) in India through Data Carpentry**

**Amit Nath**

Senior Research Fellow, Department of Library and Information Science, University of Kalyani.

## **ABSTRACT**

This research aims to measure the research impact of 27 private Higher Education Institutes (HEIs) in India through the lens of citations and social media. It applies data carpentry tools and techniques for gathering and formatting data from different databases. The primary bibliographic data was collected from Scopus, and the secondary data from four different databases (Dimensions.ai, Altmetric.com, Mendeley.com, and Unpaywall.org). The data is analyzed in terms of coverage, associations between variables (citations vs. altmetrics), and open access (dis)advantages of altmetric. The result indicates that 18.51% of publications from private HEIs are covered in altmetric, while 95.77% of publications are in Mendeley. Twitter event has the most extensive coverage among altmetric data, and Dr. D. Y. Patil Vidyapeeth (DYVPV) registered with the broadest coverage in altmetric. The coverage level of altmetric data is higher for institutions focused on multidisciplinary research than medical, technical, and management institutions. Almost positive correlations are observed between most of the variables. Moreover, the coverage of open access publications in altmetric is considerably higher than those of non-OA articles. The open access altmetric advantages (OAAA) and categorical open access altmetric advantages (COAAA) confirmed the open access advantages of selected institutions across altmetric events. The overall research result suggests that discipline- and institute-specific considerations are pivotal when evaluating institutions' productivity using altmetric.

## **Keywords**

Altmetrics, Data carpentry, Data wrangling, NIRF ranking, OpenRefine, Social media visibility, Social media metrics.

## **Introduction**

Scholarly communications are considered the preferable way of disseminating human brainchild knowledge. Traditionally, the impact of scholarly artifacts has been measured by their citation counts. Nevertheless, nowadays, the unprecedented development of ICT-enabled tools, and internet technologies, especially the invention of the social web, has opened up new possibilities and completely changed the world's communication system for the last few decades. These technological transformations and innovations significantly impacted the scholarly communication process. People use social media platforms for real-time conversation or entertainment and for accessing, storing, and

disseminating scholarly resources. In the meantime, scholarly products are increasingly being mentioned and discussed on different social media platforms. Sometimes researchers upload the preprint or post-print version of the research product into social media for better visibility and access. These social media platforms have also become an alternative way of measuring scholarly publications' impact. As a result, the term "altmetrics" was introduced to measure scholarly activities over social media. Altmetrics or alternative metrics capture various aspects of research dissemination, such as mentions in social media, blogs, news outlets, policy documents, and other online platforms. It offers a more comprehensive and timely view of research impact, taking into account the increasingly diverse ways in which research is shared and consumed in today's digital world.

The most researched question in altmetric research is the presence of altmetric data across the scientific fields. Several studies have been observed to identify the presence at different levels and the research fields that are covered mostly (Fang et al., 2020; Ortega, 2020). The current study investigates the impact research of 27 private HEIs of India through the lens of scientific and public attention. More specifically, this research highlights the performance of the said institutions in terms of coverage, OA uptakes, associations between citations and altmetric events, and finally, open access (dis)advantages for altmetric data.

### **Related literature**

This section tried to highlight only four important zones related to the study area, i) Availability of altmetric for the Indian articles, ii) Coverage of altmetric at the individual, and institutional level, iii) Relationship between citations and altmetrics, iv) Altmetric coverage of Open Access articles

### **Availability of altmetric data for Indian articles**

The coverage of altmetric data always varies widely from country to discipline. Some previous literature also discussed the coverage of Indian articles in altmetric data. For instance, Banshal et al. (2018) conducted an exploratory study based on ResearchGate. They examined Indian articles listed in Web of Science and ResearchGate for 2016 and found that 61% of WoS articles have mentioned been on ResearchGate. They also found some disciplinary variation for altmetric data in ResearchGate. A similar study by (Banshal et al., 2019b) analyzed the coverage comparison of Indian and global publications in altmetric.com and found that 28.5% of publications from India were covered in altmetric, which is 18% lower than the world data. They found that Mendeley and ResearchGate were the most popular platforms among others, particularly for scholarly articles. However, the fields like medicine and biology get much more attention than those of information science and engineering. Nath & Jana (2021) examined the coverage of altmetric data for Indian "Earth and Planetary Science (EPS)" publications and compared them with world data. They found that altmetric presence was very meager; overall, 32.7% of Indian EPS publications were covered in altmetric, while 35.75% were for world data. Another study based on the coverage in Mendeley for Indian EPS articles was observed by Nath et al. (2021) and found that 98.57% of publications have covered in Mendeley, and the presence level was varied by discipline. Banshal et al. (2019a) conducted another study on the disciplinary variation of altmetric data for Indian articles and found that disciplines like Medical, Biological, and Multidisciplinary have above 60% of publications with altmetric attention. Conversely, Mathematics, Material Sciences, and Engineering have less than 25% altmetric attention. Similarly, platform-wise variations also have been analyzed. It is found that Twitter and Mendeley have more coverage than Facebook and News. Shrivastava & Mahajan (2017) conducted a study by analyzing the coverage of publications and also established the relation among altmetric events and citation counts in the Physics discipline and found that Twitter has the highest (22.68%) coverage, followed by Facebook (3.62%), and Blogs (2.18%).

### **Coverage of Altmetric at the institutional level**

There have already different types of studies of altmetric coverage at the institutional level. Those study mainly focuses on the comparative analysis of those selected institutions' altmetric coverage and other related factors. Solanki et al. (2019) analyzed the social media coverage of the most productive Indian Institutions. They found that, on average 28.5% of publications were found in altmetric. However, the coverage percentage varies from 5-60% for different institutions. They concluded that discipline-specific institutes, like Medical and Biological Science highly mentioned on social media platforms compared to others. Lamba et al. (2021) studied 669 publications of Computer Science sub-disciplines from 35 central universities using Altmetric Explorer and Dimensions. They found that, as per attention score, Jawaharlal Nehru University (JNU) had ranked at the top, and the largest number of readers were from the University of Hyderabad. Also, Twitter received the most mentions in altmetrics, followed by Google+, patents, and finally, Facebook. Holmberg, Bowman, et al. (2019) focused on the important reason which influences the number of altmetric events surrounding the scientific articles from that institution. The result showed that to capture altmetric attention, an international connection is essential, and the relationship between the researcher and the institution must be strong. Still, on the other side, the institutional research profile is not always reflected in the altmetric counts.

### **Relationship between citations and altmetrics**

There have been lots of studies already published on the web that have measured the relationship or associations between citations and altmetrics in journal-specific (Barakat et al., 2018, 2019; Chang et al., 2019; Huang et al., 2018; Nath et al., 2020; Nocera et al., 2019); discipline-specific (Amath et al., 2017; Barbic et al., 2016; Costas et al., 2015; Mohammadi & Thelwall, 2014; Nath et al., 2021; O'Connor et al., 2017), and region-specific (Ali & Richardson, 2017; Eldakar, 2019; Gorraiz et al., 2018). A study conducted by Bornmann (2015) provides an idea on the three most popular altmetrics, i.e., Twitter (microblogging), Mendeley and CiteULike (reference managers), and Blogs. The study specifically concentrated on the correlations between altmetric and citation counts. After calculation, the correlation with citations and microblogging was small ( $r=0.003$ ) and for Blogs ( $r = 0.12$ ), CiteULike ( $r = 0.23$ ), and Mendeley ( $r = 0.51$ ). An extensive analysis was carried out by Costas et al. (2015) across scientific fields to determine the relationships between altmetric and citations. They found that the density of altmetric counts was very low and not frequent for research outputs, where 15% to 24% of articles have altmetric attention. The correlation result was positive but relatively weak. Thelwall et al. (2013) conducted an excellent comparative study taking 11 altmetric events and WoS citations, where 76 to 208,739 publications had at least one altmetric event in each case. A significant association was found among variables. Similarly, some studies have been conducted to determine the relationship between citations and altmetrics for Indian publications. Nath & Jana (2021) analyzed the correlation between citations and 11 altmetric events on narrow subject-categories of "Earth and Planetary Sciences (EPS)" articles and found a weak correlation. However, the correlation value for India and the world share was positive. Another study was done by Banshal et al. (2021) with a large dataset from WoS, specifically for Indian publications and altmetric data through ResearchGate and altmetric.com; the relationship was a weak positive correlation. Lamba et al. (2021) analyzed the data of 35 Indian central universities on Computer Science, and interestingly the relationship was positive between readers and dimensions citations.

### **Altmetric coverage of Open Access articles**

Previous studies also identified the openness of scholarly article has a significant role in getting more altmetric attention than paywall articles. For instance, Dehdarirad & Didegah (2020) took 83,444 publications (articles and reviews) on Life and Bio-medicine Science disciplines for four years (i.e., 2012-2016) listed in the Medline database. The objective was to predict the social media visibility of

the OA publications and their possible factors and found that, the coverage of OA publications mentioned was higher than those of non-OA publications. The average Tweets, Facebook posts, News, and blogs increased by 92.7%, 25.7%, 83.9%, and 48.4%, respectively. Adie (2014) studied a multidisciplinary hybrid journal named Nature Communications and found that OA articles got more social media attention on Mendeley and Twitter than non-OA articles. Schultz (2021) explored collecting the data from Altmeteric.com and WoS between 2010 and 2018 and found that Gold, Green, and Hybrid OA publications positively correlated with news mentions. After all, OA articles remained higher than those paywall articles. Holmberg, Hedman, et al. (2019) focused on the relative visibility of how often OA articles get at least one altmetric mention on the investigated altmetric platforms, how frequently they receive the mentions, and the discipline-wise variation. They found that OA publications from Economic Geography, Psychology, Social and Veterinary Sciences received more attention in citations and social media. In contrast, the OA articles from Medicine and Health sciences get fewer. The result significantly agreed on the discipline & platform-specific consideration when assessing the impact of OA articles.

The current body of literature focused on journal-, discipline-specific to determine the altmetric coverage of scholarly publications. However, no specific research incorporated altmetric analysis for private HEIs in respect of measuring the magnitude of associations as well as the advantage of the openness of articles of Indian origin. In this study, we tried to fill the research gap and measure the alternative impact of private HEIs in India through altmetric.

### **Objectives and Research Questions**

The research aims to measure the impact of selected private universities in India through the lens of public attention (social media engagement) and scientific attention (citations). Moreover, this research includes data carpentry tools (OpenRefine) and techniques to identify the presence level of altmetric and citation data, the degree of associations between them, and open access advantages of altmetric data. Based on these objectives, the following research questions have been formulated:

**RQ1:** What is the coverage level of Indian private HEIs on social media platforms? Do the institution types affect the coverage levels of social media? If yes, what are the possible reasons behind this variation?

**RQ2:** Do the altmetric counts and citation counts of Indian private HEIs related to each other? If related, which altmetric events are primarily associated with the citation counts?

**RQ3:** Does the openness of articles reflect on altmetric mentions? If yes, which path of OA impacted the most?

### **Methodology**

The methodology used in this research is acknowledged and discussed in four sub-groups, i.e., selection of institutions, data collection, data reliability & data analysis.

### **Selection of Institutions**

India has had a large educational system in terms of variety and quantity since the past. There are 1,113 universities nationwide (AISHE, 2021), including 657 publicly funded (235 Central Govt. and 422 State Govt.) and 456 privately funded (both aided and unaided). Thus, ranking and selecting prominent institutions is a risky job. There are different ranking frameworks like QS University Ranking, NIRF

ranking, World University Ranking, THE (“Times Higher Education World University Ranking”), and more.

Here we consulted the National Institutional Ranking Framework (NIRF) Ranking because of two reasons: i) The framework is developed by the Ministry of Education, the Govt. of India, exclusively for Indian HEIs; ii) It includes a variety of research institutions like universities (both public and private), research institutions, technical, management institutions, etc. The overall category of the NIRF Ranking 2022 includes 100 HEIs; among them, we found 27 privately funded institutions. We have selected those 27 top private institutions (12 multidisciplinary, 9 technical, 5 medical, and 1 management institutions) for limiting our study. These institutions are as follows- (1) Amrita Vishwa Vidyapeetham; (2) Manipal Academy of Higher Education, Manipal; (3) Vellore Institute of Technology; (4) Siksha ‘O’ Anusandhan; (5) Birla Institute of Technology & Science – Pilani; (6) Kalinga Institute of Industrial Technology; (7) SRM Institute of Science and Technology; (8) Amity University; (9) Saveetha Institute of Medical and Technical Sciences; (10) Chandigarh University; (11) Shanmugha Arts Science Technology & Research Academy; (12) Kalasalingam Academy of Research and Education; (13) Koneru Lakshmaiah Education Foundation University; (14) Thapar Institute of Engineering and Technology; (15) Lovely Professional University; (16) JSS Academy of Higher Education and Research; (17) Symbiosis International; (18) Sathyabama Institute of Science and Technology; (19) Dr. D. Y. Patil Vidyapeeth; (20) Sri Ramachandra Institute of Higher Education and Research; (21) Banasthali Vidyapith; (22) SVKM's Narsee Monjee Institute of Management Studies; (23) Datta Meghe Institute of Medical Sciences; (24) Shiv Nadar University; (25) Bharath Institute of Higher Education & Research; (26) Sri Sivasubramaniya Nadar College of Engineering; (27) University of Petroleum and Energy Studies.

## **Data Acquisition**

This research implies two types of datasets, primary data (mainly bibliographic metadata) and secondary data (citation data, altmetric data, open access data, and readership data).

### **Primary Dataset**

The primary dataset for this study was collected from the Scopus database because of its coverage (Jana et al., 2020). We searched the database through the affiliation ID search mechanism to retrieve relevant documents for each institution. The search results restricted two additional filters: i) the publication year limited to a single year (i.e., 2020) to give considerable time for getting the optimum number of citations, usually 2-3 years after publication; and ii) document types as Research Article (AR) and Review Article (RE) because of these two formats of the document are considered as first reported manifestation of brainchild knowledge. After meeting the selection criteria, the retrieved documents were downloaded in the 27 ‘CSV’ files separately and merged those files into one ‘CSV’ through the OpenRefine data wrangling tool. Finally, we found 22,553 documents that 27 private universities in India published. Among those, 19,564 (86.75%) publications have been found with active DOI (Table 1).

### **Secondary Dataset**

The entire secondary data acquisition process is based on the Digital Object Identifier (DOI). Therefore, the articles with active DOI were incorporated individually to collect further data from diverse platforms through OpenRefine. We collected the secondary data from four ODbL (Open Data Commons Open Database License) based databases (citation data from Open Citations, altmetric data from Altmetric.com, open access data from Unpaywall.org, and readership data from Mendeley.com) through REST/API calls against each DOI. These databases produce JavaScript Object Notation (JSON) formatted text as a response against each API call and are stored in OpenRefine. Furthermore,

each response has been observed, analyzed, and parsed the desired data using General Refine Expression Language (GREL). The detail methodology of data fetching was discussed in the previous research by Nath & Jana (2023). The descriptive statistics of data coverage are presented (Table 2), and the data acquisition process was done in February 2023.

### Data Analysis

The dataset was analyzed to measure the alternative impact of research publications by privately-funded HEIs in India. The descriptive statistics of research and altmetric profile were performed using Excel. We only consulted six altmetric events (Twitter, Facebook, Wikipedia, News, Blogs, and Readers) because these were mostly covered events for Indian publications, as reported by previous studies (Banshal et al., 2019b; Nath & Jana, 2021). To measure the degree of association between variables (Citations vs. Altmetric events), we performed the Spearman Correlations ( $\rho$ ) instead of Pearson due to the skewing problem of the dataset. Results were interpreted using the scale given by (Akoglu, 2018). The  $\rho$  values were calculated for altmetric and citation data with 95% confidence level for total datasets and for individual institutions.

Further, we investigated the open access altmetric advantages (OAAA) and categorical open access altmetric advantages (COAAA) based on the entire dataset. The OA Altmetric Advantages (OAAA<sub>i</sub>) can be measured for all types of altmetric events in two ways, i.e., article-based and altmetric-based. OAM represents either the mean of OA articles that are mentioned in the altmetric event (article-based) or the mean of the altmetric event for all OA articles in the altmetric event *i*. Similarly, NOAM refers to the same value for non-OA articles. The OAAA of altmetric event can be measured as:

$$OAAA_i = \frac{OAM_i - NOAM_i}{NOAM_i} \times 100$$

We proposed, the categorical open access altmetric advantages (COAAA) is used to access the OA advantages based on the OA channels. It is calculated by specific OA publications (e.g., Gold OA) with respect to other OA publications (e.g., Green, Bronze, and Hybrid). Here, SOAM refers to the mean of specific OA (Gold) articles that are mentioned in altmetric event (article-based) or the mean of the altmetric event for all specific OA articles in the altmetric event *i*, and OOAM refers to the same value for other OA (Green/Bronze/Hybrid) articles. Similarly, other OA categories have been calculated. The COAAA can be measured across altmetric events as follows:

$$COAAA_i = \frac{SOAM_i - OOAM_i}{OOAM_i} \times 100$$

The value of OAAA<sub>i</sub> = 50 means that OA publications are mentioned 50% more than non-OA publications (OA advantage) on a specific altmetric event, Twitter. On the other hand, a negative value of OAAA<sub>i</sub> indicates OA articles are mentioned less than non-OA articles (OA disadvantage). Similarly, the COAAA<sub>i</sub> value of 50 (-50) for Gold OA indicates that Gold OA publications are 50% (-50%) more (less) mentioned than other OA categories.

The data analysis and visualization were performed in R programming using ‘dplyr’, ‘ggplot2’, and ‘ggcorrplot’ packages.

Table 1: Brief Overview of Selected Institutions

Rank	Name of the Institutes	Abbreviation	Location	Year of Establishment	Total Publications	Publications with DOI
16	Amrita Vishwa Vidyapeetham	AVV-C	Coimbatore, Tamil Nadu	1994	943	857
17	Manipal Academy of Higher Education	MAHE	Manipal, Karnataka	1953	1979	1810
18	Vellore Institute of Technology	VIT-T	Vellore, Tamil Nadu	1984	2298	2117
30	Siksha 'O' Anusandhan	SOA-B	Bhubaneswar, Odisha	1996	955	884
32	Birla Institute of Technology & Science	BITS	Pilani, Rajasthan	1964	1106	1071
34	Kalinga Institute of Industrial Technology	KIIT	Bhubaneswar, Odisha	1992	669	611
36	S.R.M. Institute of Science and Technology	SIST	Chennai, Tamil Nadu	2002	1629	1356
42	Amity University	AMI-U	Gautam Budh Nagar, Uttar Pradesh	2005	1478	1277
44	Saveetha Institute of Medical and Technical Sciences	SIMT	Chennai, Tamil Nadu	2005	2238	1641
48	Chandigarh University	CHU-M	Mohali, Punjab	2012	544	298
49	Shanmugha Arts Science Technology & Research Academy	SATR	Thanjavur, Tamil Nadu	1984	709	686
50	Kalasalingam Academy of Research and Education	KARE	Srivilliputtur, Tamil Nadu	1984	356	312
54	K L College of Engineering	KLCE	Vaddeswaram, Andhra Pradesh	1980	1153	848
57	Thapar Institute of Engineering and Technology	TIET	Patiala, Punjab	1956	950	913
58	Lovely Professional University	LPU-P	Phagwara, Punjab	2005	976	801
60	JSS Academy of Higher Education and Research	JAHR	Mysuru, Karnataka	2008	427	395
62	Symbiosis International	SYM-P	Pune, Maharashtra	1971	456	379
67	Sathyabama Institute of Science and Technology	SIST	Chennai, Tamil Nadu	1987	634	530
76	Dr. D. Y. Patil Vidyapeeth	DYPV	Pune, Maharashtra	1956	142	135
83	Sri Ramachandra Institute of Higher Education and Research	SRHR	Chennai, Tamil Nadu	1985	339	318
87	Banasthali Vidyapith	BNV-B	Banasthali, Rajasthan	1935	277	247
89	SVKM's Narsee Monjee Institute of Management Studies	NMMS	Mumbai, Maharashtra	1981	256	234
92	Datta Meghe Institute of Medical Sciences	DMMS	Wardha, Maharashtra	1990	766	749
94	Shiv Nadar University	SNU-U	Dadri, Uttar Pradesh	2011	221	212
95	Bharath Institute of Higher Education & Research	BIHR	Chennai, Tamil Nadu	1984	394	294
96	Sri Sivasubramaniya Nadar College of Engineering	SNCE	Kancheepuram, Tamil Nadu	1996	335	306
97	University of Petroleum and Energy Studies	UPES	Dehradun, Uttarakhand	2003	323	283
Total					22553	19564

Sources	Methods used	Data Sources	Number of Articles	%
Altmetric data	REST/API calls through OpenRefine	Altmetric.com	3622	18.51
Readership data		Mendeley.com	18736	95.77
Citations data		Dimensions.ai	14315	73.17
Open Access data		Unpaywall.org	18709	95.63

## Results

The entire result section has been divided into three parts, i.e., research and altmetric profile of private HEIs of India, testing associations between variables (citations vs. altmetric events) at different levels (overall, individual institutions), and finally, OA (dis)advantages in altmetric.

## Research Profile

Before going into an in-depth analysis, let us introduce the quantitative research profile of selected private HEIs through the eye of scientific impact. A total of 27 private HEIs (12 multidisciplinary, 9 technical, 5 medical, and 1 management) have produced 22,553 research publications collaboratively. Among those, 19,564 publications have active DOI, produced by 34,446 faculty members (faculty/item is 0.57), and received 261,225 citations (citations/item is 13.35). The descriptive statistics of publications, citations (based on the Dimensions.ai), h-index, and faculty strength (collected from the NIRF Ranking website) were presented in Table 3. In terms of research outputs, VIT-T ranked at the top with 2,117 publications and an h-index of 68, followed by MAHE (1,810 publications and h-index of 530, SIMT (1,641 publications and h-index of 32), and SRIST (1,356 publications and h-index of 47). When we consulted the CPI value, DYPV holds the first position with a CPI of 60.52, followed by KIIT (27.53), BITS (26.16), TIET (18.88), and LPU-P (17.77). When we consider the item/faculty, SIMT ranked at the top with a value of 1.97, followed by TIET (1.61), DMMS (1.38), BITS (1.28), and SNCE (1.28).

Rank	Name of the Institutes	Total Publication	H-index	Faculty Strength	Item/Faculty	Sums of Citations	Citations/Item (CPI)
16	AVV-C	857	34	1830	0.47	6,955	8.12
17	MAHE	1810	53	2617	0.69	29,835	16.48
18	VIT-T	2117	68	2633	0.80	28,627	13.52
30	SOA-B	884	42	1161	0.76	7,926	8.97
32	BITS	1071	50	835	1.28	28,020	26.16
34	KIIT	611	40	1898	0.32	16,821	27.53
36	SRIST	1356	47	3624	0.37	14,019	10.34
42	AMI-U	1277	55	1871	0.68	22,441	17.57
44	SIMT	1641	32	831	1.97	3,827	2.33
48	CHU-M	298	28	2217	0.13	2,949	9.90
49	SATR	686	40	722	0.95	8,564	12.48
50	KARE	312	31	546	0.57	4,599	14.74
54	KLCE	848	29	1024	0.83	4,631	5.46
57	TIET	913	61	566	1.61	17,237	18.88
58	LPU-P	801	55	2371	0.34	14,235	17.77
60	JAHR	395	28	874	0.45	3,320	8.41
62	SYM-P	379	25	1258	0.30	3,012	7.95
67	SIST	530	40	1260	0.42	6,203	11.70
76	DYPV	135	18	583	0.23	8,170	60.52

83	SRHR	318	17	750	0.42	1,708	5.37
87	BNV-B	247	23	621	0.40	2,248	9.10
89	NMMS	234	24	554	0.42	2,367	10.12
92	DMMS	749	17	543	1.38	10,462	13.97
94	SNU-U	212	26	246	0.86	3,281	15.48
95	BIHR	294	23	1896	0.16	2,136	7.27
96	SNCE	306	29	268	1.14	3,693	12.07
97	UPES	283	30	847	0.33	3,939	13.92

# collected from the NIRF Ranking website

### Altmetric Profile

The altmetric and readership data for 27 private HEIs have been fetched from the Altmetric and Mendeley database, respectively. The altmetric events have been analyzed in two formats, i.e., overall and institutional presence. There are 3,622 publications (18.51%) that have been mentioned with at least one altmetric event, and 18,736 publications (95.77%) mentioned in Mendeley amongst the total. The SNU-U has become the top as 40.57% of publications have mentioned in altmetric, followed by MAHE having 38.34% of coverage and DYPV having 35.56%. Conversely, SIMT has the lowest rank in terms of altmetric attention; only 3.11% of publications are covered, followed by KLCE having 4.95% and DMMS having 4.94%. According to the altmetric event, Twitter holds the first position for getting 16.73% of publications coverage, followed by News having 2.07%. On the other side, Wikipedia has the lowest coverage having 0.67% of publications. However, the individual coverage percentage also varies across institutions. For instance, the SNU-U ranked at the top, having 39.62% of publications covered by Twitter, followed by MAHE with 35.80% of coverage. The DYPV ranked the top among the observed institutions regarding coverage in all social media platforms, except for Twitter and Mendeley (Figure 1). In Mendeley, the highest coverage comes from CHU-M, with 99.66% of publications, followed by BNV-B (99.60%), SNU-U (99.06%), and SYM-P (98.94%).

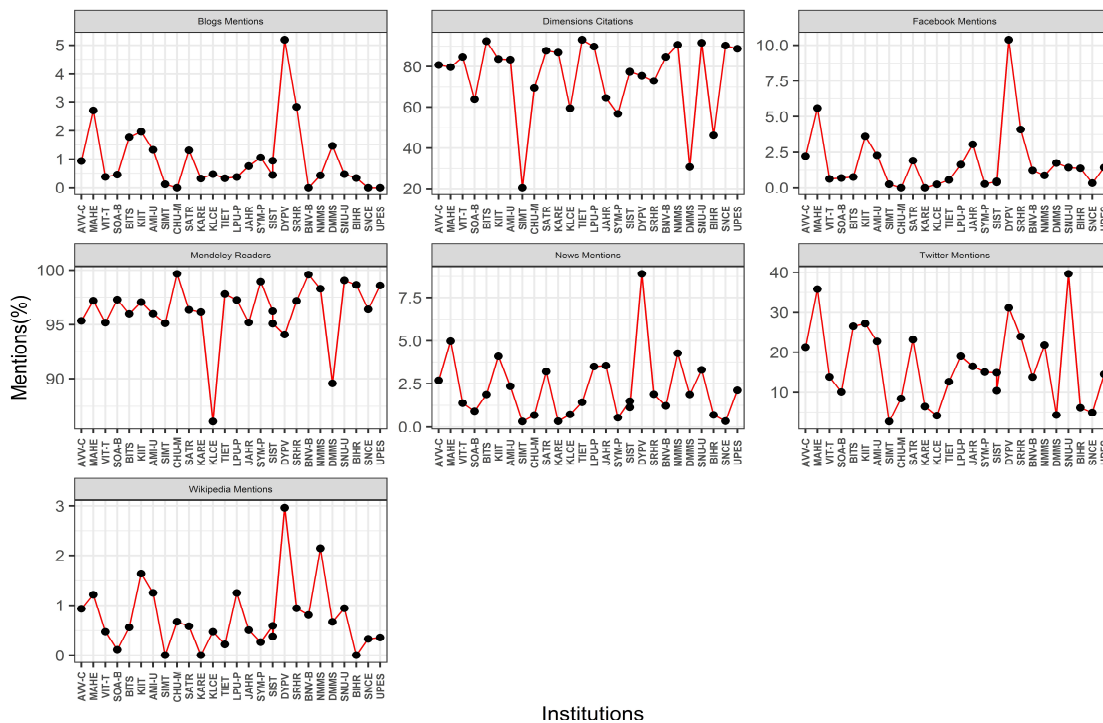


Figure 1: Institute-wise coverage of altmetric and citation data for private HEIs.

### Testing Associations between variables

A total of 3,622 publications from 27 private HEIs have been found with at least one altmetric event and incorporated to measure the degree of association among the variables, i.e., citations (Dimensions citations) and six altmetric events (Mendeley readership, Twitter, Blogs, News, Wikipedia, and Facebook). The Spearman correlation study was applied with a 95% confidence level for measuring the associations at two levels (overall and institution-wise). The overall associations between variables have been observed using the correlation matrix (Figure 2). Here, the blue-color indicates the positive rho values, whereas brown-color for negative rho value, the intensity of the color and size of the square indicates how strong the rho value is, whether positive or negative. We observed very strong correlations between DC vs. MR and a strong correlation for BM vs. NM, with rho values of 0.67 and 0.33, respectively. On the other hand, the rest of the variables are relatively low as they are negligible to weakly correlate with each other. Only TM has, on average, a moderate range of positive correlation with others except with WPM (zero rho value).

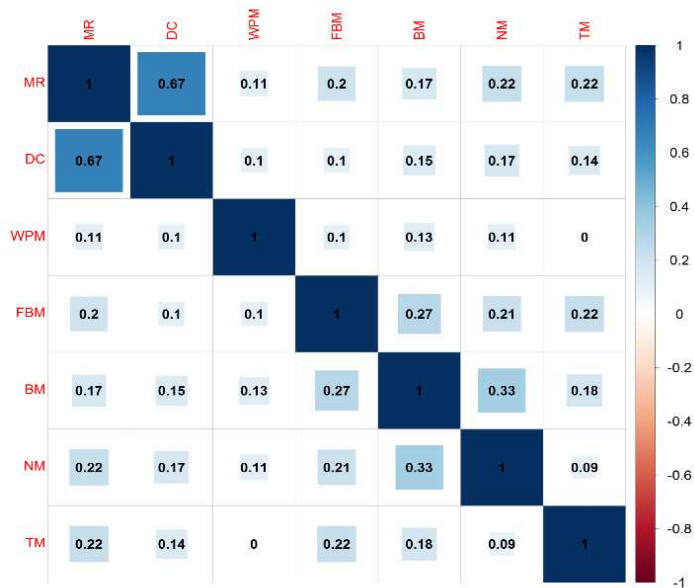


Figure 2: Correlation matrix (n= 3,622) for observed data (MR= Mendeley Readership, DC=Dimensions Citations, WPM= Wikipedia Mentions, FBM= Facebook Mentions, BM= Blog Mentions, NM= News Mentions, TM= Twitter Mentions)

At the institutional level, the correlational analysis was performed between Dimensions Citations (DC) and altmetric events. Figure 3 represents institution-wise rho values of six altmetric events, where two colors indicate different meanings. The black square-dots indicate the rho values for institutions, and the blue line indicates the mean rho value for each altmetric event. Two institutes, DYPV and DMMS, showed moderate to strong correlations across all altmetric events. Among the events, a powerful correlation has been identified for Mendeley readers with citations for every observed institution. Where the mean rho value for RS is 0.72, strong to very strong correlations have been found for all institutions with citations, highest for SNCE (0.93), followed by KARE (0.86), DMMS (0.85), TIET (0.84), DYPV (0.80), and SRHR (0.78). For Facebook, the average rho value is 0.10, ranging from -0.28 (KLCE) to 0.84 (DMMS). In the context of Blogs, the average rho value is 0.13, where the DYPV became the top with the rho value of 0.70, followed by DYPV (0.54) and also four institutions (SOA-B, SATR, SBU-U, and KARE) with negative rho values. For Wikipedia, the average rho value is 0.10, and nineteen institutions have registered weak to moderate positive correlations. While five intuitions

with negligible to weak negative correlations. The mean rho value 0.16 and 0.18 have registered for News and Twitter, respectively. Fifteen institutions have more than the mean value in News, whereas eleven institutions for Twitter. Along with positive rho values, we observed some institutions have registered with zero rho values across altmetric events, such as SIMT, KARE, and BIHR for Wikipedia, CHU-M, KARE, and TIET for Facebook, BNV-B, CHU-M, SNCE, and UPES for Blogs.

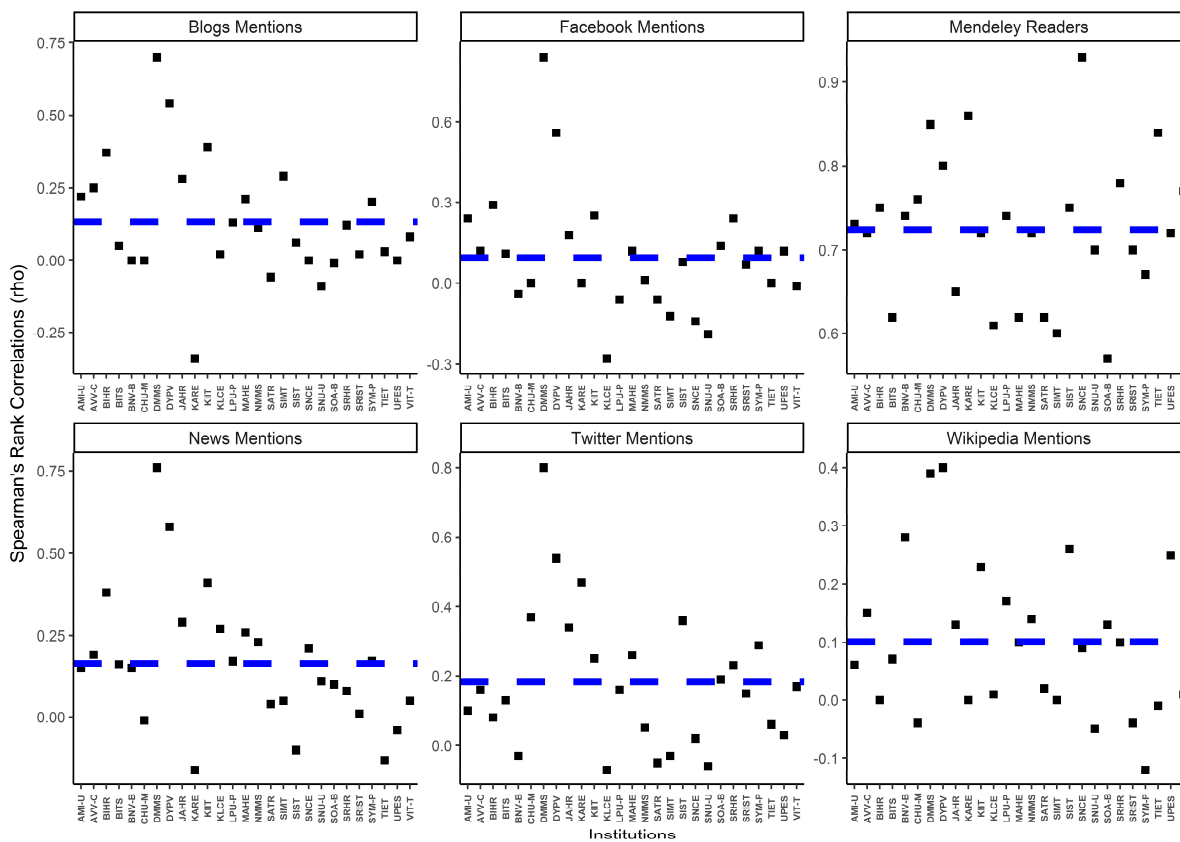


Figure 3: The rho values for six altmetric events and citations across the institutions.

### OA vs. Non-OA

In this section, we discussed whether the openness of publications has highly correlated in all altmetric events compared to close access or not. The open access status of publications has gathered been from the Unpaywall database. Among the total publications with altmetric events, we found 3,598 publications with OA status (1,806 OA and 1,792 non-OA) and incorporated them in the analysis.

Figure 4 reveals the correlation plot of OA and non-OA publications across the institutions. The upper triangle (right) plots the correlation values of OA publications, whereas non-OA publications are in the lower triangle (left). It is observed that the intensity of the blue colour box is more in the upper triangle than in the lower one. As the intensity grows, the rho values increase, and the correlation becomes strongly positive. Moreover, All the OA publications correlate more significantly with citations than non-OA articles, except for Dimensions citations. Except for Mendeley, all variables registered with negligible to weak correlations for non-OA publications. Some negligible negative correlations (DC vs. FBM, FBM vs. WPM, WPM vs. TM, and TM vs. BM) were also found for non-OA publications. Similarly, for OA publications, the rho value shows positive values for all pairs, strong correlations

between DC and MR, and other altmetric events showed weak to moderate correlations. The result indicates that OA publications are more frequently mentioned in altmetric than non-OA publications.

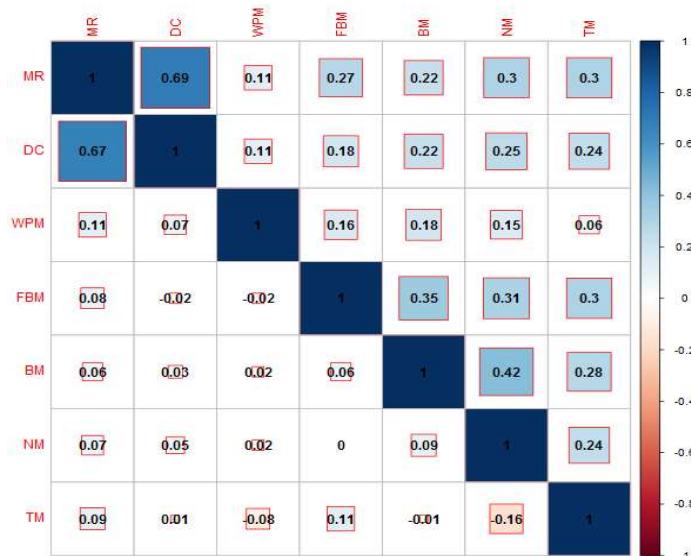


Figure 4: Correlation matrix for mentions and citations of OA (upper triangle; n= 1,806) and Non-OA articles (lower triangle; n=1,792). MR= Mendeley Readers, DC= Dimensions Citations, WPM= Wikipedia Mentions, FBM= Facebook Mentions, BM= Blog Mentions, NM= News Mentions, TM= Twitter Mentions.

### OA (dis)advantages

The OA altmetric (dis)advantages (OAAA) measure the altmetric mentions of OA publications compared to non-OA articles. There are two ways of calculating the OAAA. One is article-based (the number of OA and non-OA publications mentioned in altmetric compared to those that are not), and altmetric-based (calculated by the average event of both OA and non-OA publications that are mentioned in altmetric). The OAAA for all altmetric events has been observed and presented in Figure 5. According to the altmetric-based OAAA, the highest value was observed for NM (1093.50%), followed by TM (938.71%), BM (888.00%), and the lowest value observed for MR (78.65%) and WPM (180.51%). FBM has the lowest to moderate value of 380.12%. On the other hand, BM has the highest value of article-based OAAA of 404.97%, and the lowest value is 0.64% for MR, followed by TM (50.84%).

Similarly, the categorical OA altmetric advantages (COAAA) is introduced to measure the OA (dis)advantages according to OA paths. The value of COAAA is calculated based on a particular OA category (e.g., Green) compared to other OA categories (Figure 6). It is also calculated in two ways, article-based COAAA: the number of publications mentioned in altmetric-events in a specific category compared to other OA categories, and altmetric-based COAAA: the average number of mentions in the specific OA category concerning other OA categories. According to altmetric-based COAAA, Hybrid OA occupies an advantageous position across all altmetric events. The News mentions registered with the highest advantages (1054.58%), followed by Blogs (1039.25%), Twitter (928.81%), Facebook (423.13%), Mendeley (232.58%), and Wikipedia (165.17%). Interestingly, Gold OA (the most popular path of OA) holds the disadvantageous position for all events, while Green and Bronze OA have advantages for three (Blogs, News, and Mendeley) and four (Facebook, Twitter, Wikipedia, and

Mendeley) events, respectively. When we looked at the article-based COAAA, the Green OA holds the highest advantageous position compared to other OA paths for all altmetric events. In contrast, Gold OA is disadvantageous for all events.

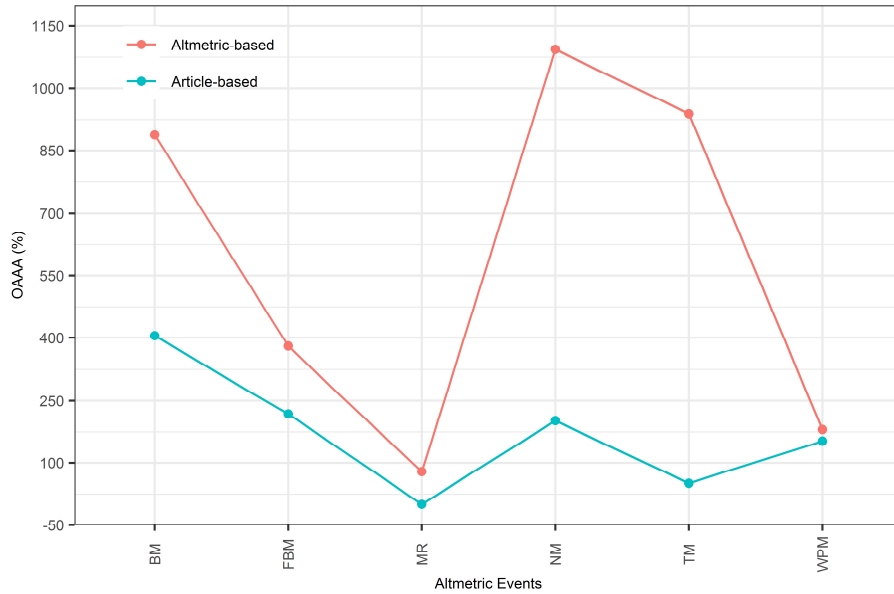


Figure 5: OA altmetric advantages of articles across altmetric events

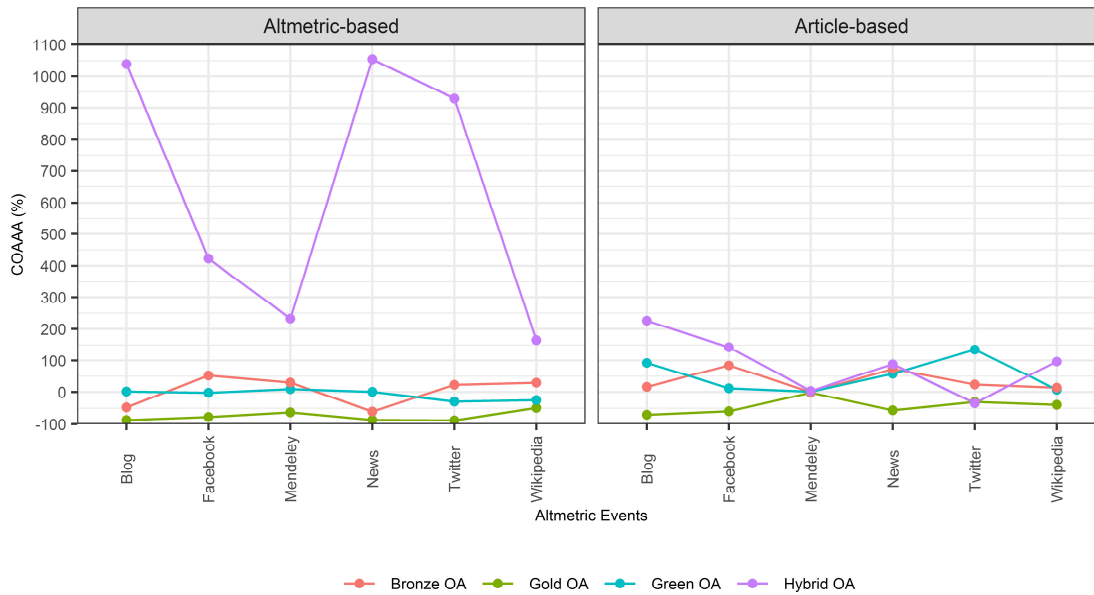


Figure 6: Categorical OA advantages of articles across altmetric events

## Discussion

This study measured the altmetric performance of India's top 27 private HEIs according to NIRF ranking 2022 (overall category). The primary dataset (mainly bibliographic metadata) was collected from the Scopus database, and secondary data from four ODbL-based databases, namely, Altmetric.com, Dimensions.ai, Unpaywall.org, and Mendeley.com. These private HEIs produced 22,553 research publications for 2020 and 19,564 publications with active DOI. The research profile of private HEIs is systematically presented. The research portfolio indicates: i) the top producing institutions (VIT-T, MAHE, SIMT, and SIST) are mainly from the southern part (Tamil Nadu, Karnataka) of India. Among the selected HEIs, 13 institutions were located in the southern states of India, six from the western states, four from the northern states, and two from central and eastern states of India. It indicates an enthusiastic environment of higher education in the southern parts of India; ii) the institutes with large research outputs do not necessarily receive the highest CPI values compared to those with small research outputs. In contrast, VIT-T, MAHE, SIMT, SIST, and AMI-U are the top five research-producing institutions with CPI values of 13.52, 16.48, 2.33, 10.34, and 17.57, respectively. Whereas, DYPV has made only 135 research publications and received the highest CPI of 60.52, due to skewing distribution of citations, two articles got 4,150 and 2,230 citations, respectively; iii) the institutes with large research outputs do not necessarily mean the faculty strength is higher or vice-versa. For instance, SIMT's total research output and items per faculty are 1,641 and 1.97, respectively. In contrast, VIT-T has a total research output, and items per faculty are 2,117 and 0.80, respectively, which is relatively low items per faculty compared to SIMT. Similarly, with only 306 total research output, SNCE has good items per faculty of 1.14; iv) we also noticed a strong collaboration network between five institutes (MAHE, KIIT, DYPV, AMI-U, and DMMS). These institutions have published several highly cited articles collaboratively.

The altmetric portfolio of 27 private HEIs was also analyzed and presented. Based on the total publications, we found that 18.51% of publications are covered in altmetric, whereas 95.77% of publications are covered in Mendeley. The coverage of altmetric data is still inadequate, increasing over the years as found by previous studies for Indian publications (Banshal et al., 2019b; Nath et al., 2021; Nath & Jana, 2021; Solanki et al., 2019). Some previous studies also observed altmetric data coverage issues at a global scale (Costas et al., 2015; Fang et al., 2020; Ortega, 2020). According to presence level, Mendeley has the largest, followed by Twitter, News, Facebook, Blogs, and Wikipedia. The coverage level of altmetric data is higher for institutions focused on multidisciplinary research than medical, technical, and management institutions.

To answer the RQ2, we performed a correlational study between variables at different levels. The results confirmed strong correlations between DC vs. MR and a strong correlation for BM vs. NM; the rest of the pairs are moderate to weak correlations. Several previous studies also confirmed the strong rho values for Mendeley readership and Scopus/WoS/Dimensions citations (Eldakar, 2019; Lamba et al., 2021; Mohammadi & Thelwall, 2014; Nath et al., 2021). We found zero correlations for the overall category between TM and WPM because only a few publications have been mentioned in WPM. The institution-wise correlation between readership and citations showed surprising results (an almost perfect correlation is found for SNCE). We found higher rho values for all investigated institutions compared to previous studies (Eldakar, 2019; Mohammadi & Thelwall, 2014). To identify the open access benefits of altmetric data, we calculated OAAA and COAAA. The open access status of publications collected from the Unpaywall database. Our findings are similar to previous studies (Adie, 2014; Dehdarirad & Didegah, 2020; Holmberg, Hedman, et al., 2019), and the OAAA values also confirmed that open access articles are significantly mentioned in altmetric compared to non-OA articles. It is noticed that the value of altmetric-based OAAA is more than the value of article-based OAAA across the altmetric events because of the mentioned density (the number of post counts is higher) of publications.

## Conclusion

This research aims to measure the impact of research publications made by 27 private HEIs of India through the lens of scientific and public attention. A total of 19,564 publications have found been with active DOI; only 18.51% of publications are covered in altmetric and 95.77% of publications in Mendeley. VIT-T has ranked the top as per publications, while DYPV holds the first position as per the altmetric coverage. The openness of publications plays a momentous role to be mentioned in altmetric. Based on the data sources and methodology used in this research can be extended for measuring the alternative impact of individual entities (like journals, institutions, authors, and countries) and also comparing with others.

Besides the findings described above, however, this research also has some methodological limitations. Firstly, several databases are available in the bibliographic domain in both ODbL-based (OpenAlex, Semantic Scholar, Crossref, MAG, Dimensions.ai, and more) and subscription-based (WoS) that provides metadata against the search query. Here, we collected the bibliographic metadata from Scopus, leading to bias. Secondly, we collected citation data from the Dimensions database, which may lead to bias because citation counts mostly depend on the coverage of the database. Thirdly, we applied Spearman's correlations method to test the associations. However, it does not provide any causal relationship between variables, but it is the easiest way to see how variables are related (Huang et al., 2018; Sud & Thelwall, 2014; Thelwall et al., 2013). Lastly, we only selected the institutions listed in the overall category of NIRF 2022; a mass number of institutions are avoided for this reason. Therefore, results may vary when other institutions are included.

## References

- All India Survey on Higher Education 2020-21* (Govt. Report No. 11th; pp. 1–310). (2021). Ministry of Education, Govt. of India.
- Adie, E. (2014). *Attention! A study of open access vs non-open access articles*. <https://doi.org/10.6084/m9.figshare.1213690.v1>
- Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, 18(3), 91–93. <https://doi.org/10.1016/j.tjem.2018.08.001>
- Ali, M. Y., & Richardson, J. (2017). Pakistani LIS scholars' altmetrics in ResearchGate. *Program*, 51(2), 152–169. <https://doi.org/10.1108/PROG-07-2016-0052>
- Amath, A., Ambacher, K., Leddy, J. J., Wood, T. J., & Ramnanan, C. J. (2017). Comparing alternative and traditional dissemination metrics in medical education. *Medical Education*, 51(9), 935–941. <https://doi.org/10.1111/medu.13359>
- Banshal, S. K., Singh, V. K., Kaderye, G., Muhuri, P. K., & Sánchez, B. P. (2018). An altmetric analysis of scholarly articles from India. *Journal of Intelligent and Fuzzy Systems*, 34(5), 3111–3118. <https://doi.org/10.3233/JIFS-169495>
- Banshal, S. K., Singh, V. K., & Muhuri, P. K. (2021). Can altmetric mentions predict later citations? A test of validity on data from ResearchGate and three social media platforms. *Online Information Review*, ahead-of-print(ahead-of-print). <https://doi.org/10.1108/OIR-11-2019-0364>
- Banshal, S. K., Singh, V. K., Muhuri, P. K., & Mayr, P. (2019a). Disciplinary variations in altmetric coverage of scholarly articles. In Catalano G., Daraio C., Gregori M., Moed H.F., & Ruocco G. (Eds.), *Int. Conf. Scientometrics Informetrics, ISSI - Proc.* (Vol. 2, pp. 1870–1881).
- Banshal, S. K., Singh, V. K., Muhuri, P. K., & Mayr, P. (2019b). How much research output from India gets social media attention? *Current Science*, 117(5), 753–760. <https://doi.org/10.18520/cs/v117/i5/753-760>
- Barakat, A. F., Nimri, N., Shokr, M., Mahtta, D., Mansoor, H., Masri, A., & Elgendy, I. Y. (2019). Correlation of Altmetric Attention Score and Citations for High-Impact General Medicine Journals: A Cross-sectional Study. *Journal of General Internal Medicine*, 34(6), 825–827. Scopus. <https://doi.org/10.1007/s11606-019-04838-6>

- Barakat, A. F., Nimri, N., Shokr, M., Mahtta, D., Mansoor, H., Mojadidi, M. K., Mahmoud, A. N., Senussi, M., Masri, A., & Elgendy, I. Y. (2018). Correlation of Altmetric Attention Score With Article Citations in Cardiovascular Research. *Journal of the American College of Cardiology*, 72(8), 952–953. <https://doi.org/10.1016/j.jacc.2018.05.062>
- Barbic, D., Tubman, M., Lam, H., & Barbic, S. (2016). An Analysis of Altmetrics in Emergency Medicine. *Academic Emergency Medicine*, 23(3), 251–265. <https://doi.org/10.1111/acem.12898>
- Bornmann, L. (2015). Alternative metrics in scientometrics: A meta-analysis of research into three altmetrics. *Scientometrics*, 103(3), 1123–1144. <https://doi.org/10.1007/s11192-015-1565-y>
- Chang, J., Desai, N., & Gosain, A. (2019). Correlation Between Altmetric Score and Citations in Pediatric Surgery Core Journals. *Journal of Surgical Research*, 243, 52–58. <https://doi.org/10.1016/j.jss.2019.05.010>
- Costas, R., Zahedi, Z., & Wouters, P. (2015). Do “altmetrics” correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science and Technology*, 66(10), 2003–2019. <https://doi.org/10.1002/asi.23309>
- Dehdarirad, T., & Didegah, F. (2020). To What Extent Does the Open Access Status of Articles Predict Their Social Media Visibility? A Case Study of Life Sciences and Biomedicine. *Journal of Altmetrics*, 3(1), <https://doi.org/10.29024/joa.29>
- Eldakar, M. A. M. (2019). Who reads international Egyptian academic articles? An altmetrics analysis of Mendeley readership categories. *Scientometrics*, 121(1), 105–135. <https://doi.org/10.1007/s11192-019-03189-7>
- Fang, Z., Costas, R., Tian, W., Wang, X., & Wouters, P. (2020). An extensive analysis of the presence of altmetric data for Web of Science publications across subject fields and research topics. *Scientometrics*. <https://doi.org/10.1007/s11192-020-03564-9>
- Gorraiz, J., Blahous, B., & Wieland, M. (2018). Monitoring the broader impact of the journal publication output on country level: A case study for Austria. In Erdt M., Theng Y., Sesagiri Raamkumar A., & Rasmussen E. (Eds.), *Commun. Comput. Info. Sci.* (Vol. 856, p. 62). Springer Verlag; [https://doi.org/10.1007/978-981-13-1053-9\\_4](https://doi.org/10.1007/978-981-13-1053-9_4)
- Holmberg, K., Bowman, T., Didegah, F., & Lehtimäki, J. (2019). The Relationship Between Institutional Factors, Citation and Altmetric Counts of Publications from Finnish Universities. *Journal of Altmetrics*, 2(1), <https://doi.org/10.29024/joa.20>
- Holmberg, K., Hedman, J., Bowman, T. D., Didegah, F., & Laakso, M. (2019). Do articles in open access journals have more frequent altmetric activity than articles in subscription-based journals? An investigation of the research output of Finnish universities. *Scientometrics*. <https://doi.org/10.1007/s11192-019-03301-x>
- Huang, W., Wang, P., & Wu, Q. (2018). A correlation comparison between Altmetric Attention Scores and citations for six PLOS journals. *PLoS ONE*, 13(4). <https://doi.org/10.1371/journal.pone.0194962>
- Jana, S., Gupta, A., & Nath, A. (2020). Assessment of Global Performance on COVID-19 Research during 1990-2019: An Exploratory Scientometric Analysis. *Qualitative and Quantitative Methods in Libraries*, 9(Special Issue on COVID19), 21–39.
- Lamba, M., Kashyap, N., & Madhusudhan, M. (2021). Research evaluation of computer science publications using Altmetrics: A cohort study of Indian Central Universities. *Global Knowledge, Memory and Communication*, ahead-of-print(ahead-of-print). <https://doi.org/10.1108/GKMC-07-2020-0097>
- Mohammadi, E., & Thelwall, M. (2014). Mendeley readership altmetrics for the social sciences and humanities: Research evaluation and knowledge flows. *Journal of the American Association for Information Science and Technology (JASIST)*, 65(8), 1627–1638. <https://doi.org/10.1002/asi.23071>

- Nath, A., & Jana, S. (2021). An Altmetric Analysis of Scholarly Publications from Earth and Planetary Science Discipline: An Exploratory Study of Indian Publications. In *Annual Proceedings of the Science & Technology Metrics* (3rd ed., pp. 138–145). Digital Information Research Labs. <https://doi.org/10.6025/stm/2021/3/138-145>
- Nath, A., Jana, S., Kerketta, S., & Kerketta, S. (2020). Who Reads PLOS Research Articles? Extensive Analysis of the Mendeley Readership Categories of PLOS Journals. *Journal of Scientometric Research*, 9(3), 245–252. <https://doi.org/10.5530/jscires.9.3.32>
- Nath, A., Jana, S., & Santra, P. P. (2021). Characteristics of Mendeley Readership for Earth and Planetary Science Articles. *DESIDOC Journal of Library & Information Technology*, 41(6), <https://doi.org/10.14429/djlit.41.6.16961>
- Nath, A., & Jana, S. (2023). Data Wrangling from Socio-Academic Web-Space: Designing a Meta Model. *Journal of Information and Knowledge*, 60(2), 113–125. <https://doi.org/10.17821/srels/2023/v60i2/170971>
- Nocera, A. P., Boyd, C. J., Boudreau, H., Hakim, O., & Rais-Bahrami, S. (2019). Examining the Correlation Between Altmetric Score and Citations in the Urology Literature. *Urology*. <https://doi.org/10.1016/j.urology.2019.09.014>
- O'Connor, E. M., Nason, G. J., O'Kelly, F., Manecksha, R. P., & Loeb, S. (2017). Newsworthiness vs scientific impact: Are the most highly cited urology papers the most widely disseminated in the media? *BJU International*, 120(3), 441–454. <https://doi.org/10.1111/bju.13881>
- Ortega, J.-L. (2020). Altmetrics data providers: A meta-analysis review of the coverage of metrics and publication. *Profesional de La Información*, 29(1), <https://doi.org/10.3145/epi.2020.ene.07>
- Schultz, T. (2021). All the research that's fit to print: Open access and the news media. *Quantitative Science Studies*, 2(3), 828–844. [https://doi.org/10.1162/qss\\_a\\_00139](https://doi.org/10.1162/qss_a_00139)
- Shrivastava, R., & Mahajan, P. (2017). An altmetric analysis of ResearchGate profiles of physics researchers: A study of University of Delhi (India). *Performance Measurement and Metrics*, 18(1), 52–66. <https://doi.org/10.1108/PMM-07-2016-0033>
- Solanki, T., Karmakar, M., Banshal, S., & Singh, V. (2019). Social Media Coverage of Research Output from 100 Most Productive Institutions in India. *Journal of Scientometric Research*, 8(3), 143–149. <https://doi.org/10.5530/jscires.8.3.30>
- Sud, P., & Thelwall, M. (2014). Evaluating altmetrics. *Scientometrics*, 98(2), 1131–1143. <https://doi.org/10.1007/s11192-013-1117-2>
- Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, C. R. (2013). Do Altmetrics Work? Twitter and Ten Other Social Web Services. *PLoS ONE*, 8(5), e64841. <https://doi.org/10.1371/journal.pone.0064841>