



Science Mapping of Open-Access Articles on Data Visualisation: A Scopus-based Analysis

Dr. Iranna M. Shettar

Librarian (I/c), Central Library, National Institute of Technology Warangal - 506 004, Telangana State, INDIA

ARTICLE INFO

Key words: Data Visualization, Open Access, Science Mapping, Scopus Database, Publishing Trends

doi: 10.48165/lt.2024.10.1

ABSTRACT

The prime objective of the present study is to analyze the impact of open-access publishing on research publications in data visualization journal research papers using scientometric indicators and science mapping techniques. A total of 9,018 publication records were extracted from the Scopus database for five years, from 2018 to 2022. The parameters used for the present study are open-access publication trends and their impact on the number of publications and citations, most prolific authors, organizations, countries, journals, etc. The study results indicate that 60.55% of articles on data visualization were published as open access, and the trend of open-access publication has grown exponentially. The study further found that open-access articles have a better academic impact than overall articles published. Huamin Qu of Hong Kong University of Science and Technology has contributed the highest number (22) of open-access articles on data visualization; the authors affiliated with the Centre National de la Recherche Scientifique (CNRS), Paris, have contributed the highest (116) OA articles; and the United States was the topmost productive country with the highest (1,809) OA articles. IEEE Transactions on Visualization and Computer Graphics was the most preferred journal among researchers on data visualization to publish their research as open-access articles.

Introduction

Data visualization presents data and information in a graphical context using visual elements. Data visualization tools provide an easy-to-understand and easy-to-use means to see and understand data trends, outliers, and patterns via visual components such as charts, graphs, and maps. The primary objective of data

visualization is to communicate information clearly and efficiently to users.

According to IBM (2024), data visualization is “the representation of data through the use of common graphics, such as charts, plots, infographics, and even animations.” Netquest (2021) defines “data visualization as the process of acquiring, interpreting, and comparing data in order to

^{*}Corresponding author.

E-mail address: imshettar@gmail.com (Iranna M. Shettar)

Received 25.09.2024; Accepted 06.10.2024

Copyright @<https://acspublisher.com/journals/index.php/lt/>

clearly communicate complex ideas, thereby facilitating the identification and analysis of meaningful patterns.”

In simple words, data visualization is the process of using visual elements to represent data. Data visualization helps people to understand and work with large and complex amounts of data. Data visualization presents complex data to the common people and particular audiences without any technical knowledge about the data or information.

Data visualization encompasses various techniques and tools to translate complex data sets into understandable and actionable insights. Today, data visualization plays a crucial role in multiple fields, including business, science, government, and journalism, helping decision-makers and analysts make sense of vast amounts of data and communicate their findings effectively.

Open-access publishing is an essential reason for increased research and research publications in scientific and academic communities and helps to promote the research and its ‘citability.’ Citability is the “ability of a publication to be citable.” Many factors determine a publication’s ‘citability,’ and these factors may differ from subject to subject and geography to geography. Citation is a scientific impact of a research publication and is used as an indicator to evaluate its scientific value (Aksnes et al., 2019). The current study focuses on the open-access factor that plays a vital role in determining the citability of research publications.

Literature Review

Previously, various studies have been presented evaluating the impact of open-access publications as a means to gain more academic and research citations in different subject domains. Content that is easily and freely accessible has a better chance of reaching a wider audience and gaining more global visibility, which leads to increased reading and may result in more possibilities for citation (Harnad & Brody, 2004; Kousha & Abdoli, 2010; Wang et al., 2015; Holmberg et al., 2019; Shettar & Hadagali, 2020; Shettar, 2022).

Hardly any metric studies are being conducted to analyze the research publications on ‘Data Visualization.’ However, a research publication was found that presented a scientometric analysis of the Information Visualization research publications.

Singh, Ranjan, and Rai (2019) presented a scientometric study using the Information Visualization Research Publications indexed in the Scopus database for the period between 1990 and 2018. The study ana-

lyzed 6,192 research publications published in nine different forms, and the most preferred research publication form was a Journal Article, with more than 90%. The study found a constant increase in yearly publications and observed multiple authorship trends. The United States was the most prolific country, and most authors were affiliated with the Centre National de la Recherche Scientifique (CNRS). Nucleic Acids Research is the most preferred journal, and the National Institutes of Health was the highest funding agency for publishing research on information visualization.

Significance & Limitations of the Study

The present study will help analyze the role of open-access publishing in achieving a higher academic and research impact in the data visualization research literature. The study has potential significance for researchers in the field of data visualization in designing their research priorities and publishing policies.

It is also essential to study the open-access publication priorities compared with closed-access and impact comparisons according to global trends in data visualization research publications. Hence, this study is limited to evaluating the research impact of open-access publishing in the data visualization journal research papers published between 2018 and 2022, according to the Scopus database.

Objective of the Study

The main objective of the current study is to evaluate the impact of open-access publishing on research publications in the data visualisation journal research papers published between 2018 and 2022. The other objectives are:

- To study the global publications and open-access trends in journal publications;

- To evaluate the academic and research impact of open-access journal publications;

- To identify highly prolific authors, organizations, and countries involved in publishing open-access journal research papers;

- To find the most preferred journals among researchers in data visualisation.

Materials and Methods

The bibliographic records of the journal research articles for the study were extracted from the Scopus database, pub-

lished between 2018 and 2022. Of the 9,018 indexed journal paper records on data visualisation research, 5,460 journal papers published as open access were further extracted for in-depth analysis in Excel format. The extracted data were further analysed using MS Excel and statistical formulae. Science network maps were developed using VOSviewer to analyse the collaborations and co-authorships between the authors, organisations, and countries.

Results and Discussions

Year-wise distribution of journal articles

Table 1 presents the analysis of articles published in “Data Visualisation” in various journals between 2018 and 2022. The table indicates the exponential growth in the number of articles during the study period; the lowest number of research articles published in 2018 (1,199) increased year on year, and the highest number of research articles published in 2022 (2,562). However, the citations received for these publications do not follow any trend, with the highest number of citations received for the articles published in 2019, followed by 2020 and 2018. The lowest number of citations was received for the latest publications in 2022, which clearly indicates the publications attract more visibility with every passing year, which is further proved with the help of average citations per paper calculated, which clearly demonstrates that the older the publications, the higher the citations. The highest ACPP is calculated for publications in 2018 and the lowest in 2022.

Similarly, the h-index, a quality indicator to evaluate research impact, shows the highest h-index for the publications in 2018 (78) and the lowest in 2022 (41). Overall, during the study period, 9,018 articles on “Data Visualisation” have received 144,013 citations at the rate of 15.97 citations per paper and an h-index of 128. Figure 1 presents the publications and citation charts.

Table 1: Year-wise distribution of articles on Data Visualisation

Year	TP	TC	ACPP	H-Index
2018	1,199	30,524	25.46	78
2019	1,568	36,322	23.16	76
2020	1,675	34,503	20.60	75
2021	2,014	27,684	13.75	58
2022	2,562	14,980	5.85	41
Total	9,018	1,44,013	15.97	128

TP: Total Publications; TC: Total Citations; ACPP: Average Citations per Paper

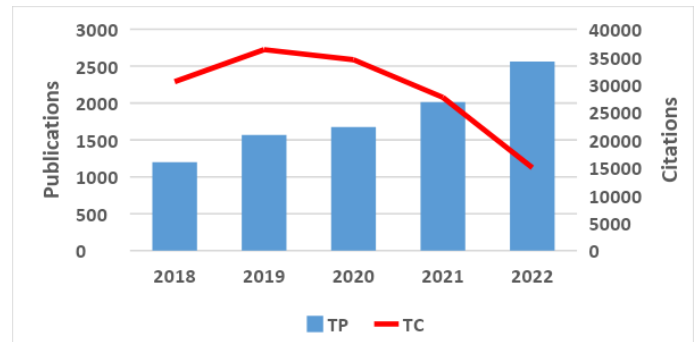


Fig. 1: Year-wise distribution of global publications and citations

Year-wise distribution of open-access journal articles

Table 2 presents the analysis of open-access articles published on “Data Visualisation” in various journals between 2018 and 2022. A total of 60.55% of articles among the overall articles were published as Open Access, and year-on-year growth in the percentage has seen exponential growth, with 49.21% open-access publications in 2018 to 64.21% in 2022. Even the number of open-access publications on “Data Visualisation” has grown exponentially. These 60.55% of open-access publications have received 68.59% of overall citations, which indicates that open-access publications can create more academic impact than closed or subscription-based access articles. The highest number of citations was received for the open-access publications in 2020, followed by publications in 2019, and the lowest number of citations by the open-access articles published in 2022. These open-access articles have received citations at the rate of 18.09 average citations per paper (ACPP), with the highest 30.76 ACPP recorded for the publications in 2018 and the lowest 6.27 ACPP in 2022.

Similarly, the highest h-index was recorded in 2018, and the lowest h-index was 39 for the publications in 2022. Figure 2 presents a comparative chart of open-access articles and citation trends. Figure 3 illustrates the comparison of the percentage of overall publications and citations received.

Table 2: Year-wise distribution of open-access articles on Data Visualisation

Year	OA-TP	% TP	OA-TC	% TC	OA-ACPP	OA-h-Index
2018	590	49.21	18,147	59.45	30.76	65
2019	897	57.21	24,663	67.90	27.49	64
2020	1,053	62.87	25,137	72.85	23.87	63
2021	1,275	63.31	20,520	74.12	16.09	53
2022	1,645	64.21	10,311	68.83	6.27	39
Total	5,460	60.55	98,778	68.59	18.09	114

OA-TP: Total open-access publications; % TP: Percentage of open-access publications among total publications; OA-TC: Total citations received by open-access

publications; % TC= Percentage of citations received by open-access publications among total publications; OA-ACPP: Average Citations per open-access paper.

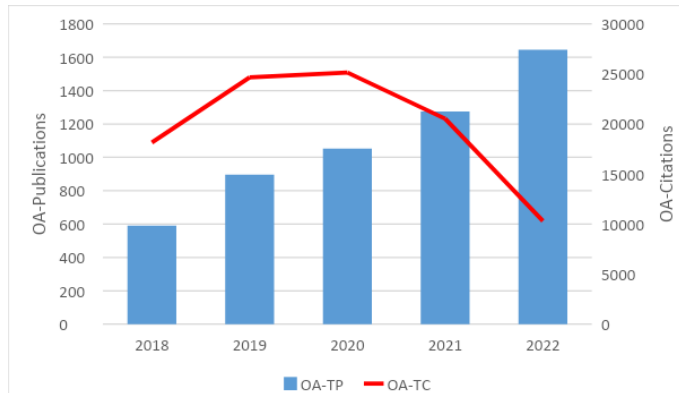


Fig. 2: Year-wise distribution of open-access publications and citations

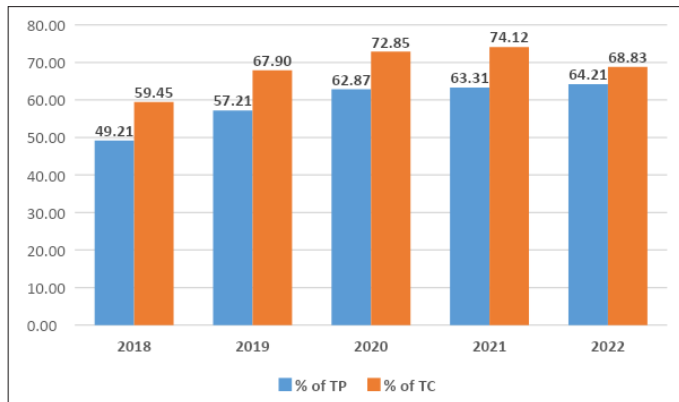


Fig. 3: Year-wise distribution of percentage of open-access publications and citations among the global publications and citations

Year-wise calculated impact of open-access journal articles

Table 3 presents an analysis of open access publications' impact compared with the overall publications on 'Data Visualization'. The study uses the analysis suggested by Shettar and Hadagali (2020), taking the year-wise average citations received by the open-access publications and overall journal articles on 'Data Visualization'. The study clearly shows the negative exponential growth; it grows slower the less time period is provided. The highest average citations per open-access article were recorded in 2018 (30.76) and the lowest in 2022 (6.27); similarly, the highest ACP was recorded for overall journal articles in 2018 (25.46) and the lowest in 2022 (5.85). The impact of open access was calculated by comparing year-wise ACP for open access and overall articles. If the computed impact value is positive and higher than 0, it will indicate a good impact, or if the calculated impact value is negative and less than 0, it will display a bad impact. The highest impact value was recorded for the publications in 2018 (5.3) and the lowest

in 2022 (0.42). The average impact value across the period was 2.12; the impact value was positive across the study period. This indicates that the open-access articles have a better academic impact than the overall publications.

Table 3: Impact of open access publications

Year	OA-ACPP	ACPP	OA-Impact
2018	30.76	25.46	5.3
2019	27.49	23.16	4.33
2020	23.87	20.6	3.27
2021	16.09	13.75	2.34
2022	6.27	5.85	0.42
	18.09	15.97	2.12

OA-ACPP: Average Citations per open-access paper; ACP: Average Citations per Paper; OA-Impact: Open-access impact calculated.

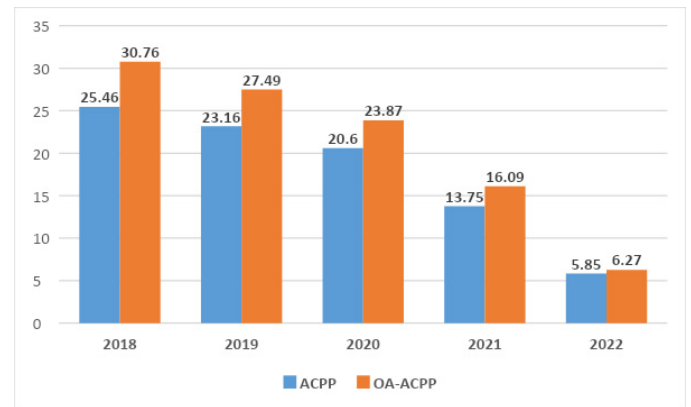


Fig. 3: Year-wise distribution of average citations per paper for global and open-access publications.

Top 10 Most Prolific Authors

Table 4 presents the list of the top 10 most prolific authors involved in the journal article publications on Data Visualization. A total of 20,757 authors were involved in publishing 5,460 open-access journal articles at the rate of 3.8 authors per paper. Huamin Qu of Hong Kong University of Science and Technology, Hong Kong, China, has published the highest number of 22 open-access research articles under study, followed by Hanspeter Pfister (17) and Wei Chen (16). Hanspeter Pfister of Harvard University, USA, has received the highest number of 1,462 citations among the top 10 authors, followed by Huamin Qu (519) and Tim Dwyer (460). Hanspeter Pfister has achieved the highest average of 86.00 citations per paper, and Huamin Qu has achieved the highest h-index of 14 among the top 10 authors. The top 10 authors have published 2.58% of articles under study and received 3.83% of citations. Figure 4 presents the co-authorship map of highly productive 500 authors, where 426 authors have collaborated with each other.

Table 4: Top 10 highly prolific authors

Author Name	Affiliation	TP	TC	ACPP	h-Index
Huamin Qu	Hong Kong University of Science and Technology, Hong Kong, China	22	519	23.59	14
Hanspeter Pfister	Harvard University, USA	17	1462	86.00	11
Wei Chen	Zhejiang University, China	16	391	24.44	10
Christin Henein	University College London, London, UK	14	7	0.50	2
Rashmi G. Mathew	University College London, London, UK	14	7	0.50	2
Kwan-Liu Ma	University of California, Davis, South Korea	13	241	18.54	8
Tim Dwyer	Monash University, Clayton, Australia	12	460	38.33	9
Benjamin Bach	University of Edinburgh, United Kingdom	11	255	23.18	8
Sheelagh Carpendale	Simon Fraser University, Canada	11	257	23.36	7
Tobias Isenberg	Université Paris-Saclay, CNRS, Inria, LISN, France	11	180	16.36	7

TP: Total Publications; TC: Total Citations; ACPP: Average Citations per Paper

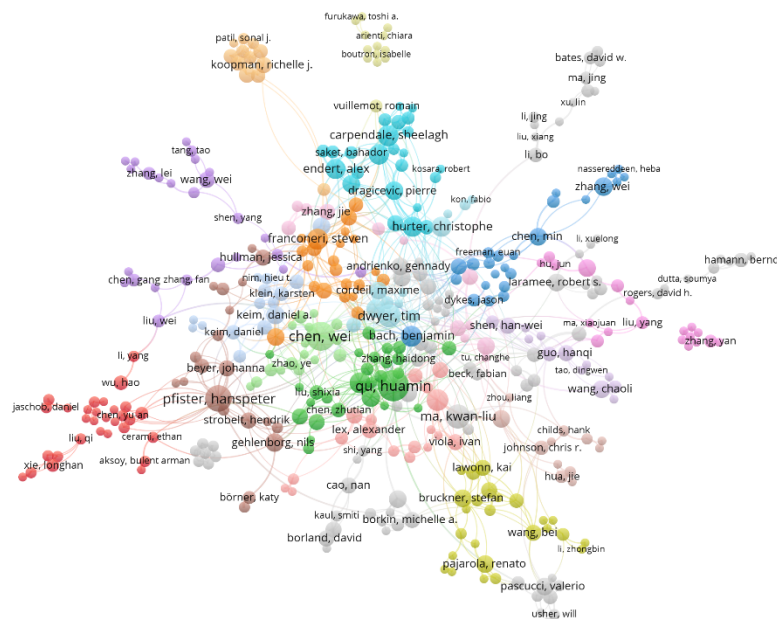


Fig. 5: Co-authorship map of highly productive authors

Top 10 Most Productive Organisations

Table 5 lists the top 10 highly productive organisations involved in journal article publications on Data Visualisation. A total of 15,536 organisations were engaged in publishing 5,460 open-access journal articles at the rate of 2.84 organisations involved per paper. Centre National de la Recherche Scientifique (CNRS), Paris, France, has published the highest 116 articles, followed by the Chinese Academy of Sciences, Beijing (97), and Harvard Medical School, Boston (68). Similarly, all

these top three organisations were ranked as the first three organisations according to the citations received and h-index. However, according to average citations per paper, the University of Toronto, Toronto, has achieved the highest 48.64 CPP, followed by Harvard Medical School, Boston, United States (41.03), and Harvard University, Cambridge, United States (38.73). These top ten organisations have published 11.74% of the articles under study and received 21.55% of the citations. Figure 4 presents a co-authorship map of highly productive organisations, where 189 have collaborated.

Table 5: Top-10 most productive organisations

Affiliation	City	Country	TP	TC	ACPP	h-Index
Centre national de la recherche scientifique (CNRS)	Paris	France	116	3118	26.88	29
Chinese Academy of Sciences	Beijing	China	97	2828	29.15	25
Harvard Medical School	Boston	United States	68	2790	41.03	23
Monash University	Melbourne	Australia	57	1367	23.98	23
University College London	London	United Kingdom	55	1874	34.07	16
Harvard University	Cambridge	United States	52	2014	38.73	20
University of Toronto	Toronto	Canada	50	2432	48.64	16
Université Paris-Saclay	Paris	France	50	1640	32.80	18
Imperial College London	London	United Kingdom	49	1493	30.47	14
University of Washington	Washington	United States	47	1727	36.74	15

TP: Total Publications; TC: Total Citations; ACP: Average Citations per Paper



Fig. 6: Co-authorship map of highly productive organisations.

Top 10 Most Productive Countries

Table 6 presents the list of the top 10 highly productive countries involved in journal article publications on Data Visualisation. A total of 128 countries were involved in publishing 5,460 open-access journal articles. The United States has the highest number with 1,809 articles, followed by China (925) and the United Kingdom (634). According to the citations received, the United States has the highest with 41,144 citations,

followed by the United Kingdom (19,941) and China (19,462). Among the top 10 productive countries, Canada has achieved the highest with 34.36 citations per paper, followed by the United Kingdom (31.45) and Italy (26.92). According to the h-index, the United States leads the table with an h-index of 83, followed by the United Kingdom (59) and China (56). Figure 5 presents a co-authorship map of the highly productive 100 countries, where 98 organisations have collaborated with each other.

Table 6: Top-10 most productive countries

Country	TP	TC	ACPP	h-Index
United States	1809	41144	22.74	83
China	925	19462	21.04	56
United Kingdom	634	19941	31.45	59
Germany	509	12508	24.57	54
Canada	281	9655	34.36	39
France	281	7218	25.69	44
Spain	249	4502	18.08	31
Australia	242	5947	24.57	40
Italy	209	5627	26.92	36
Netherlands	200	4404	22.02	33

TP: Total Publications; TC: Total Citations; ACPP: Average Citations per Paper

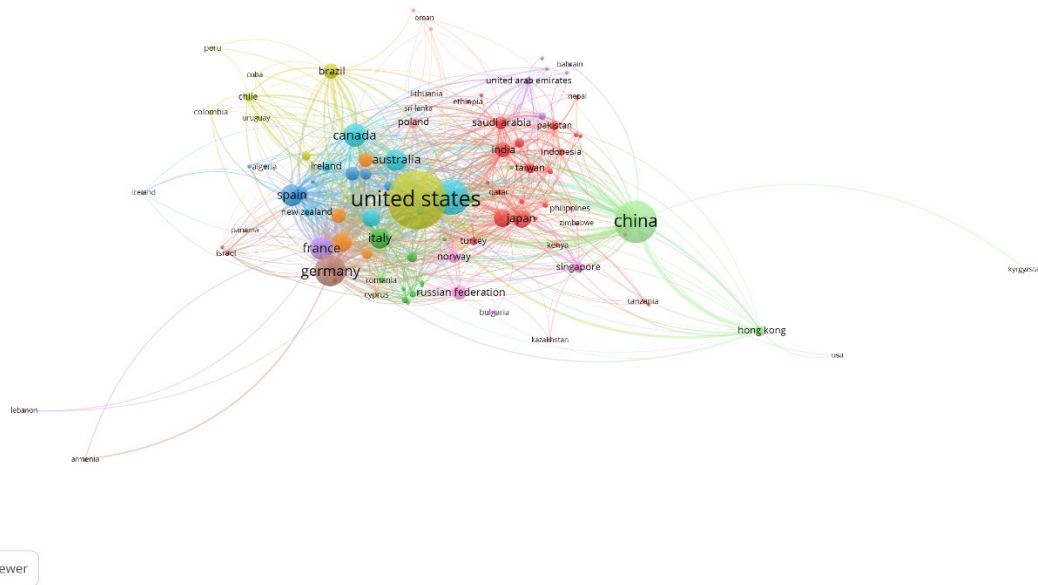


Fig. 7: Co-authorship map of highly productive countries

Top 10 Highly Preferred Journals

Table 7 lists the top 10 highly preferred journals to publish articles on Data Visualization. These 5,460 articles were published in 1,803 journals. Among the most preferred top ten journals, IEEE Transactions on Visualization and Computer Graphics has published the highest number of 336 articles, followed by IEEE Access (228) and BMC Bioinformatics (101). The journal IEEE Transactions on Visualization and Computer Graphics also received the highest number of 7,799 citations, followed by Nucleic Acids Research (4,847) and IEEE Access (3,490). The ranking of the highly preferred

journals by researchers according to average citations per paper was led by Nucleic Acids Research, which achieved 103.13 CPP, followed by Sensors-Switzerland (23.52) and IEEE Transactions on Visualization and Computer Graphics (23.21). IEEE Transactions leads the list according to the h-index on Visualization and Computer Graphics with an h-index of 46, followed by IEEE Access (33) and Nucleic Acids Research (23). The IEEE Society published the top two highly preferred journals, and two more journals were published by MDPI. The journal Nucleic Acids Research has achieved the highest cite score of 32.3 provided by Scopus based on its overall citation performance.

Table 7: Top-10 highly preferred journals

Journal Title	Cite Score (2022)	Publisher	TP	TC	ACPP	h-Index
IEEE Transactions on Visualization and Computer Graphics	10.5	IEEE Society	336	7799	23.21	46
IEEE Access	9.0	IEEE Society	228	3490	15.31	33
BMC Bioinformatics	6.2	BioMed Central	101	1406	13.92	19
Computer Graphics Forum	5.3	Wiley-Blackwell	69	1376	19.94	21
Sensors Switzerland	6.8	MDPI	65	1529	23.52	20
Remote Sensing	7.9	MDPI	64	1345	21.02	21
Visual Informatics	4.8	Elsevier	50	739	14.78	14
Plos One	6.0	PLoS	49	427	8.71	11
Nucleic Acids Research	32.3	Oxford University Press	47	4847	103.13	23
Scientific Visualisation	1.6	MEPhi	46	108	2.35	5

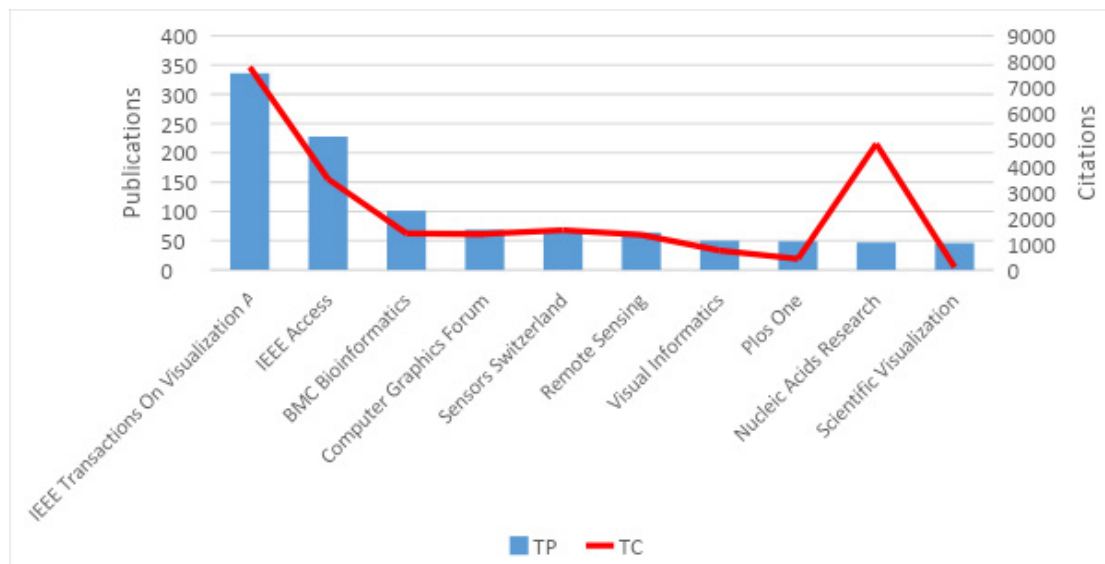


Fig. 8: Journal-wise distribution of global publications and citations

Conclusion

The present study attempts to analyze the citability of open-access publications in data visualization research literature. Data visualization has become a core research method to interpret complex data into meaningful patterns. Hence, the research publications on data visualization are increasing day by day. Usually, open-access publishing is believed to provide an extra edge to reach more readers; however, its academic and research impact may be judged only by the metric study.

The present study provides evidence that open-access publications have additional benefits in reaching a higher audience and converting readership to citation. The calculated research impact varies from 0.42 to 5.3 across the study period, and an average calculated research impact of

2.12, using the average citations per paper for both open and closed-access publications, indicates open access as a critical parameter to promote research publications. Thus, researchers in the field of data visualization need more awareness about the benefits of publishing as open access, and policymakers, administrators, and research funding agencies should support researchers in publishing their research as open access through an open-access policy.

References

- HHarnad, S., & Brody, T. (2004). Comparing the impact of open access (OA) vs. non-OA articles in the same journals. **D-Lib Magazine*, 10*(6). <https://doi.org/10.1045/june2004-harnad>

- Holmberg, K. (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. *SpringerLink*. https://link.springer.com/article/10.1007/s11192-019-03301-x?error=cookies_not_supported&code=e50092aa-0406-499a-89d5-ad32352d443e
- IBM. (2024). What is data visualisation? <https://www.ibm.com/topics/data-visualization>
- Kousha, K., & Abdoli, M. (2010). The citation impact of open access agricultural research. *Online Information Review, 34*(5), 772–785. <https://doi.org/10.1108/14684521011084618>
- Netquest. (2021). *Visualise it!: A comprehensive guide to data visualization*. <https://www.netquest.com/hubfs/docs/ebook-data-visualization-EN.pdf>
- Shettar, I. M., & Hadagali, G. S. (2021). Impact of open access publication in veterinary sciences in India: A scientometric study. In K. N. Kandapal et al. (Eds.), *Libraries: From clay tablet to fablet* (pp. 144–157). Agri-Biovet Press.
- Shettar, I. M. (2022). Increased citations through open access publishing or international collaboration: A study of engineering publications. In B. Ramesh Babu et al. (Eds.), *6th I-LISS International Conference on revitalising the libraries to the Android society* (pp. 708-715). BS Publications.
- Singh, K., Ranjan, A., & Rai, S. (2019). Information visualisation research publications during 1990-2018: A scientometric analysis. *Library Philosophy and Practice (E-journal)*. <https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=7117>
- Wang, X., Liu, C., Mao, W., & Fang, Z. (2015). The open access advantage considering citation, article usage, and social media attention. *Scientometrics, 103*(2), 555–564. <https://doi.org/10.1007/s11192-015-1547-0>