Publication status: Preprint has been submitted for publication in journal

Recalibrating the Scope of Scholarly Publishing: A Modest Step in a Vast Decolonization Process

Saurabh Khanna , Jon Ball, Juan Pablo Alperin, John Willinsky

https://doi.org/10.1590/SciELOPreprints.4729

Submitted on: 2022-09-10
Posted on: 2022-09-12 (version 1)
(YYYY-MM-DD)

Recalibrating the Scope of Scholarly Publishing:

A Modest Step in a Vast Decolonization Process

Saurabh Khanna, Stanford University

0000-0002-9346-4896

Jon Ball, Stanford University

0000-0003-2881-0016

Juan Pablo Alperin, Simon Fraser University

0000-0002-9344-7439

John Willinsky, Stanford University

0000-0001-6192-8687

Abstract

By analyzing 25,671 journals largely absent from journal counts and indexes, this study demonstrates that scholarly communication is more of a global endeavor than is commonly credited. These journals, employing the open source publishing platform Open Journal Systems (OJS), have published 5.8 million items and represent 136 countries, with 79.9 percent publishing in the Global South and 84.2 percent following the OA diamond model (charging neither reader nor author). More than half (54.6 percent) of the journals operate in more than one language, while publishing research in 60 languages (led by English, Indonesian, Spanish, and Portuguese). The journals are distributed across the social sciences (45.9 percent), STEM (40.3 percent), and the humanities (13.8 percent). For all their geographic, linguistic, and disciplinary diversity,

the Web of Science indexes 1.2 percent of the journals and Scopus 5.7 percent. On the other hand, Cabells Predatory Reports includes 1.0 percent of the journals, while Beall lists 1.4 percent of them as predatory. A recognition of the expanded scope and scale of scholarly publishing will help ensure that humankind takes full advantage of what is increasingly a global research enterprise.

Keywords: Scholarly communication, journal publishing, open access, Global South, OA diamond journals, decolonial process

Author Contributions

- Saurabh Khanna: Conceptualization, Data curation, Methodology, Formal analysis, Investigation, Visualization, Software, Resources, Writing – original draft, Writing – review & editing
- Jon Ball: Conceptualization, Data curation, Methodology, Formal analysis, Investigation, Visualization, Resources, Writing – original draft, Writing – review & editing
- Juan Pablo Alpern: Conceptualization, Methodology, Investigation, Writing – review & editing, Resources
- John Willinsky, Conceptualization, Supervision, Funding acquisition, Writing – original draft, Writing – review & editing

Recalibrating the Scope of Scholarly Publishing: A Modest Step in a Vast Decolonization Process

In 2018, Philip G. Altbach and Hans de Wit, two leading scholars of higher education, published “Too Much Academic Research Is Being Published” in *University World News*. In making their case, Altbach and de Wit point out that although “no one knows how many scientific journals there are... several estimates point to around 30,000” (2018).[1] Finding the number excessive, they declare “a crisis in academic publishing” involving “too much pressure on top journals” and “the rise of predatory journals and publishers that publish low or marginal quality research.” They recommend steps be taken to reduce the amount of research published.[2] The analysis presented below not only challenges such journal estimates but calls for a recognition of the Global South’s research commitment and engagement. The source of this analysis is 25,671 journals that are using the open source editorial management and publishing platform Open Journals Systems (OJS), first released by the Public Knowledge Project (PKP) in 2002.[3]

[1] Just how dated this number may be is suggested by the Library of Congress study of 1963 which found a global total of 35,000 journals, with 100 titles from Indonesia (Gottschalk & Desmond, 1963).

[2] Also in 2018, Gianfranco Pacchioni, Vice-Rector for Research at the University of Milano Bicocca, published *The Overproduction of Truth* (2018) on a similar theme (“accompanied by objective data and findings”) claiming that as a result of the internet, “in a short time the world of research has changed from the passionate activity of a few selected people to a crowded universe of practitioners, often with few ideas and sharing little or no ethical values” (p. 4).

[3] Founded in 1998, PKP (https://pkp.sfu.ca/) is a research and development initiative at Simon Fraser University and Stanford University.

As authors, we acknowledge in our Competing Interests note at the end of this paper our association with the Public Knowledge Project the steps taken in light of that.

Earlier studies identified 9,828 JUOJS in 2015 (Alperin, Stranack, Garnett, 2016), while a survey of 2,114 staff at these journals found 97.8 percent were open access (OA), with 13.6 percent utilizing Article Processing Charges (APC) to provide open access (Alperin, Stranack & Hanson, 2017; see also Edgar & Willinsky, 2010). These results suggest that 84.2 percent of the JUOJS offer “OA diamond” (which neither charge authors nor readers for access). A more recent study has found JUOJS constituting 60 percent of OA diamond journals among “at least 17,000, but likely up to 29,000, OA diamond journals” (Bosman et al., 2021, p. 93). Of particular relevance to this paper, Bosman and coauthors conclude that “even though they are well embedded in academic structures, OA diamond journals struggle to be properly integrated into the ecosystem of scholarly publications” (p. 84). In considering the reasons for this struggle, Altbach’s earlier work proves helpful.

Some four decades ago, in the “University as Center and Periphery," Altbach wrote of a postcolonial legacy in which the “peripheral universities" of the “Third World” are “basically distributors of knowledge... dependent on the central institutions for innovation and for direction” (1981). Their “special problems with developing research capabilities” include “few outlets for scholarship, creative writing, and research reports by Third World intellectuals” as “publishers are largely uninterested in Third World authors.”[4] Altbach saw this powerful "colonial educational heritage” a form of

[4] In other contexts, Altbach addresses “the role and nurturing of journals in the Third World” by observing that “journals are especially crucial for African scholarly and scientific development” (1998, 2), just as he

"neocolonialism" not easily overcome: "The odds against a fully autonomous and effective publishing enterprise being developed in the [African] continent are high" (1976, p. 461). The universities there, he further warns, "not only have to confront the reality of the historically and economically based power of the industrialized nations but also… the widespread desire by the industrial nations to maintain their dominant positions" (1981).[5]

The extent of this "intellectual imperialism" – as the center-periphery pattern has also been framed, beginning in the late 1960s by Syed Hussien Alatas at the University of Malaya (2000) – has also given rise to an "academic dependency" theory of relevance to this paper.[6] Its leading scholar, Syed Farid Alatas, at the National University of Singapore and the University of Malaya, writes of the multiple dependencies afflicting Global South researchers, including a "dependency on recognition" that leads to "many scholars [being] torn between satisfying the requirements of publishing in high ranked 'international' journals, particularly those listed in the [Web of Science], and publishing locally in their own languages" (2022, p. 20). The JUOJS attest to a certain break with

holds that universities in the Global South could "play a key national role in terms of training and sometimes in terms of applied research" (1981).

[5] In 2019, Collyer, Connell, Maia, & Morell observed that "the periphery continues to be a rich source of raw materials for the mainstream knowledge economy in our time. It produces data for the new biology, pharmaceuticals, astronomy, social science, linguistics, archaeology, and more. It is, for instance, a key source of data for modern climate science…. the metropole continues to be the main site of theoretical processing in the global economy of knowledge" (Collyer, Connell, Maia, & Morell, 2019, p. 10).

[6] Syed Farid Alatas on "academic dependency": "A condition in which the knowledge production of certain scholarly communities are conditioned by the development and growth of knowledge of other scholarly communities to which the former is subjected" (Alatas, 2022, p. 18).

the recognition dependency, just as these journals appear to address what the Kenyan novelist and essayist Ngũgĩ wa Thiong'o identified as a "need to move the center from its assumed location in the West to a multiplicity of spheres in all cultures of the world," which Thiong'o framed as part of a "vast decolonization process" (1991, p. xiv, 3).

In the first decade of the twenty-first century, Keynan information science scholar Ezra Ondari-Okemwa noted how the region's shortfalls in "the technological capability to support electronic knowledge transfer and scholarly publishing" were matched by regional publications' "lack [of] visibility" (2007). Similarly, Alperin and Rozemblum (2017) describe the decade between 1995 and 2015 as one of consolidation for Latin American publishing marked by efforts to gain visibility.[7]

What this paper seeks to bring to light is how editors, publishers, and researchers across the world, but especially in the Global South, have taken advantage of an open source publishing platform to create, against Altbach's odds, "an effective scholarly publishing enterprise." In the remainder of this paper, the extent of this enterprise will be mapped across the principal dimensions of scholarly publishing, grouped as follows: 1. Country and region, 2. Language, 3. Discipline, and 4. Indexing and citation. In doing so, the paper demonstrates the gains made in digital-era scholarly publishing capabilities, while also making clear that the research's visibility in the leading scholarly indexes has not grown apace. By explicitly counting thousands of journals that have too often been

[7] Allsion Muddit, CEO of the non-profit OA publisher PLOS, notes how "the need to achieve credibility and visibility within the global research system" means researchers in the Global South making a choice "between a desire to strengthen local platforms and outlets that better serve local needs and feeling pulled to 'play the game' in which norms have been set by the Global North" (2020).

discounted, we show how scholarly publishing from the Global South is shifting the assumed center to “a multiplicity of spheres.”

1.0 Data and Sample Selection

This study relies on a dataset we previously made public of 25,671 journals that published five or more items in 2020 using OJS (Khanna et al., 2021).[8] Those journals were identified through OJS’ optional PKP Beacon. Available since 2015, the beacon is used by PKP to notify OJS users of security patches and other software upgrades.[9] The journal identifier sent to PKP by the beacon contains a link for retrieving journal indexing data (using the Open Archives Initiative Metadata Harvesting Protocol). The data collected includes the journal’s title, ISSN, number of items published, titles, abstracts, and other indexing information. Only journals with five or more items published in 2020 were included (following the Directory of Open Access Journals standard for “active” journals; DOAJ, 2020). This amounted to 25,671 journals (36.5 percent of the 70,214 OJS beacons). The active journals average 38.1 items in 2020, for a total that year of 996,000 items, while these journals have published 5.4 million items since their inception. The journals are distributed across 9,793 OJS installations, with a mean of 2.62 journals per installation.

[8] The entire public dataset consists of 70,214 observations of PKP software installations, which includes 487 instances of Open Monograph Press and 100 of Open Preprint Systems, in addition to 69,627 journals using OJS.

[9] One measure of the proportion of journals using OJS that have an active beacon is provided by the 4,326 journals in the Directory of Open Access Journals (DOAJ) using OJS, for which it could be determined that 75.7 percent transmit data to PKP through the beacon.

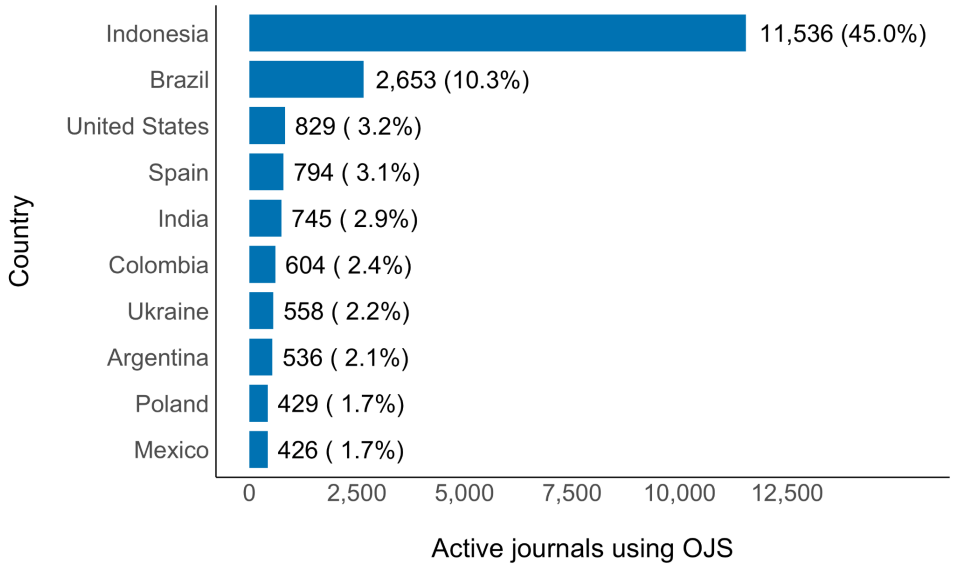
2.0 Country and Region

2.1 Country and Region Methodology

A journal's country of origin of a journal through a three step process. First, we used the ISSN (International Standard Serial Number) of a journal to query the LOC ISSN lookup API for location, which provides us with a MARC code. This MARC code can then be mapped to a country (Raoni, 2021). MARC codes that map to within-country regions or states are handled manually to reflect corresponding countries. Second, we extract the top level domains from the journal URLs and map them to their corresponding ISO 3166-1 alpha-2 country codes. Third, we geolocated the IP address of the journal host using a weekly updated copy of the GeoLite2 Free Geolocation Data.[10]

Figure 1. The distribution of active journals using OJS in top 10 countries (*N* = 19,110).

[10] Notwithstanding the robustness of our three level checks, this method can generate results skewed towards the location of hosting servers and data centers, which in itself can reduce visibility of journals originating in the Global South (Cloudscene, 2022).



Note

1. Jupyter Notebook analysis and R visualization.

2.2 Country and Region Results

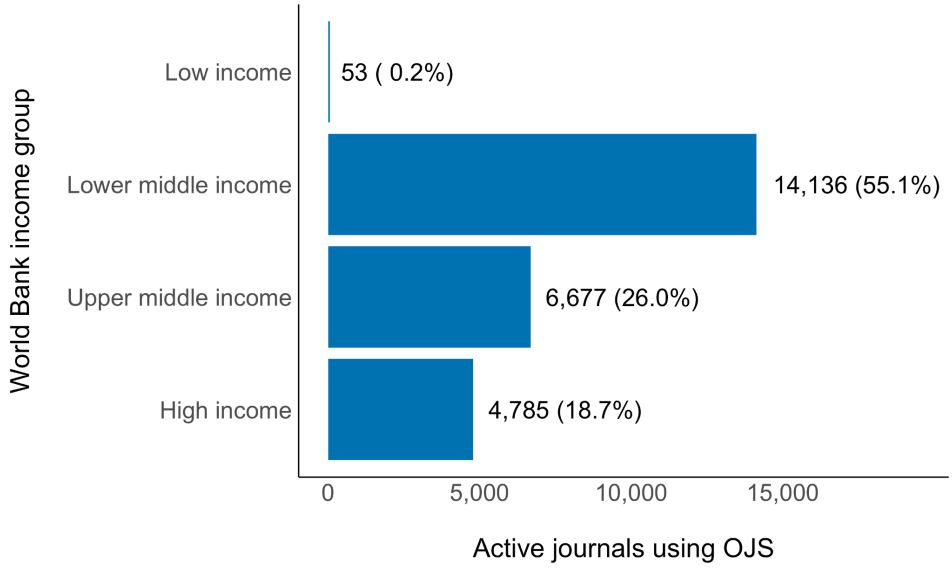
Based on the beacon data, we were able to identify 136 countries in which 25,651 (99.9%) JUOJS are published (with 20 journals for which a country of publication could not be identified). Indonesia leads with 11,535 (45.0%) journals, followed by Brazil with 2,653 (10.3%).[11] The top ten countries account for 74.6 percent of the total (Figure 1).[12] The economic distribution of JUOJS was analyzed using the four World Bank income

[11] Indonesian authors have been identified by Heather Piwowar in a 2017 study as open access leaders with 81 percent of the articles they've authored being open, compared to a world average of 41 percent and the next closest country Colombia at 64% (Van Noorden, 2019).

[12] With regard to the 558 JUOJS in Ukraine (out of an estimated total of 2,500 titles in the country; Hawkins, 2022), and the multinational science academies statement "Action Steps for Rebuilding Ukraine's Science, Research, and Innovation (2022)," PKP has reached out to the journals' editors with strategies for journal preservation involving the Internet Archive, as well the PKP Preservation Network.

groups (Figure 2). By gross national income (GNI) per capita, 81.1 percent of the JUOJS are found in the World Bank's middle income countries, with the 27 low income countries accounting for only 53 JUOJS (0.2%). This amounts to 81.6 percent of the journals coming from countries associated, by national income, with the Global South (Haug, 2021). From a geopolitical perspective, using the five United Nations Regional Groups, 78.1 percent of the JUOJS are located in three UN regional groups associated with Global South (Dados & Connell, 2012): Asia-Pacific States (54.7%), Latin American and Caribbean States (21%), and African States (2.4%). Either way, roughly 80 percent of the JUOJS are published in the Global South.

Figure 2. The distribution of journals using OJS in World Bank income groups, which are based on gross national income (GNI) per capita (*N* = 25,651).



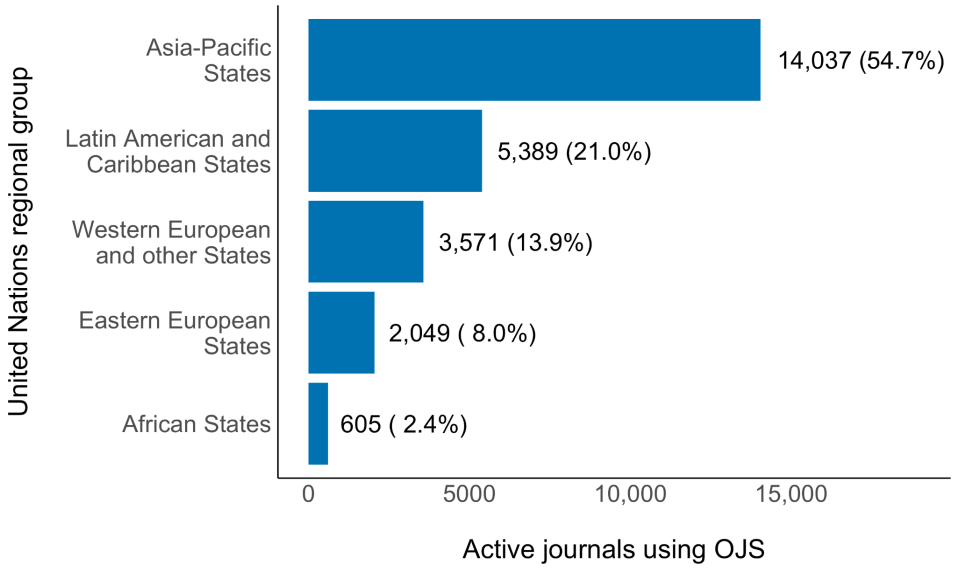
Notes

1. Jupyter Notebook analysis and R visualization.

2. Source: *World Bank Country and Lending Groups* (2021).

In comparison, the global distribution of journals in Elseveir's journal database Scopus does not include Indonesia among its countries with more than 200 titles, while only the United States appears in Scopus' and the JUOJS' top ten countries (STM, 2021, p. 15). Equally striking is the contrast in the coverage of the Global South, with Scopus including 1,085 journals (6.1 percent) from China and India among the top ten countries represented in its database, compared to the JUOJS' 16,498 journals (64.3 percent) from the Global South among its top ten. For more on the relationship between JUOJS and Scopus, see the section below on indexing.

Figure 3. Journals by United Nations regional group (*N* = 25,651)



Note

1. R analysis and visualization.

3.0 Language

3.1 Language Methodology

To determine the languages of publication for the JUOJS, we ran Google's Compact Language Detector v3 (gcld3) on the 100 most recent articles in the ISSN-verified subset of JUOJS. Gcld3 is a freely available pre-trained neural classifier and has a built-in flag for language predictions that indicates when the predicted language classification exceeds an optimal probability threshold. Only reliable language classifications were retained for each journal. Then, the most frequent language classification for each journal was designated as the journal's primary language of publishing. If gcld3 classified a given journal's article abstracts as being in multiple languages (i.e., with at least 5 articles per language), then the journal was designated as multilingual. Finally, a variety of heuristic checks were applied to verify the primary language of publishing for each journal, for example, by checking predicted languages against top-level domains, which resulted in the semi-manual correction of 478 journals. Two relatively under-resourced languages, Balochi and Faroese, are currently unsupported by gcld3 and were instead discovered by searching top-level domains.

3.2 Language Results

Using language detection strategies led to the identification of 60 languages in which research is being published in JUOJS, with a journal identified for each (Table 1). For all ISSN-verified JUOJS (*n* = 22,561), language classifications were shared with colleagues

at Google Scholar to facilitate the indexing of journals in languages not previously indexed for research work.[13]

Table 1. The 60 languages in which abstracts and articles are published in journals using OJS (*N* = 25,671).

Afrikaans	Danish	Igbo	Russian
Albanian	Dutch	Italian	Scottish Gaelic
Arabic	English	Japanese	Serbian
Armenian	Estonian	Kazakh	Sinhala
Indonesian	Faroese	Kiswahili	Slovak
Malay	Filipino	Korean	Slovenian
Balochi	Finnish	Kurdish	Spanish
Basque	French	Lithuanian	Swedish
Belarusian	Galician	Macedonian	Tamil
Bosnian	Georgian	Nepali	Thai
Bulgarian	German	Norwegian	Turkish
Catalan	Greek	Persian	Ukrainian
Chinese	Hindi	Polish	Urdu

[13] This resulted, for example, in the Google Scholar indexing of the journals *Hanken* and *Balochistaniyat*, which publish in Balochi (an Iranic language spoken in Pakistan, Afghanistan, and Iran). On the other hand, *Fróðskaparrit*, which may be the only journal to publish research in Faroese, has yet to be indexed. As discussed below in the section on indexing, Google Scholar includes the vast majority of JUOJS, including those publishing in Kazakh, Kiswahili, and Kurdish.

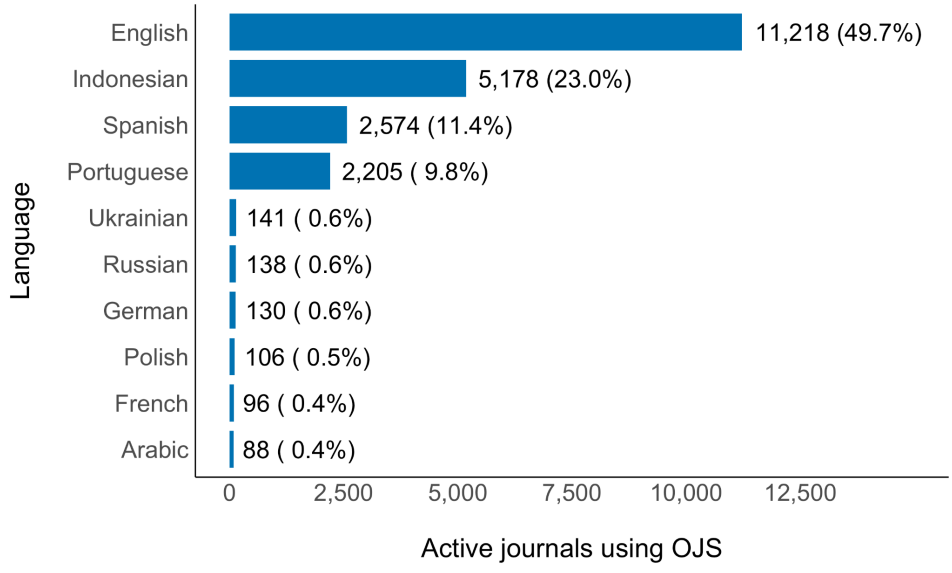
Croatian	Hungarian	Portuguese	Uzbek
Czech	Icelandic	Romanian	Vietnamese

Note

1. An example of a JUOJS that publishes in each of these languages can be found here.

Close to half of the JUOJS published articles in English (49.7 percent), with the top ten languages accounting for 97.0 percent of the titles (Figure 4). By comparison, Scopus has publications in 40 languages, with the proportion of documents (rather than journals) in English at 92.6%, followed by 2.8% in Chinese and 1.3% percent in Spanish (Vera-Baceta, Thelwall & Kousha, 2019).

Figure 4. Top ten primary languages of publication for active journals using OJS (*N* = 22,561).



Note

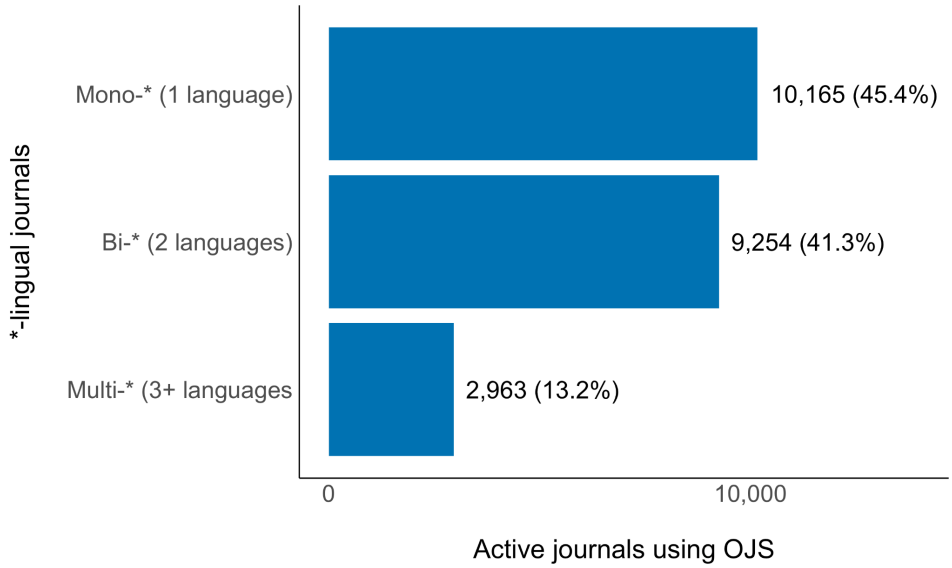
1. Jupyter Notebook analysis and R visualization.

The high level of linguistic diversity among JUOJS calls into question the common linguistic assumption that “English is the language of science,” as a recent headline in *Nature* read (Elnathan, 2021). The reign of English in research and scholarship only began to take hold in the latter half of the twentieth-century, after sharing the spotlight earlier with French and German; and while medieval Latin served as the language of learning in the West, the influx of learning from the Islamic world in Arabic, through the twelfth-century translation movement, made all the difference (Hunter-Konos, 2015, Willinsky, 2017, pp. 117-152).

Today, English hegemony is facing challenges in the name of “bibliodiversity,” which Monica Berger, New York City College of Technology, places at the center of decolonizing the open access movement in research (2021; Shearer, 2019). A “position statement” on this question of research and language has been issued by the Chilean scholar Federico Navarro, La Universidad de O'Higgins, and eleven other faculty from around the globe which would “challenge assumptions made about the use of English as a ‘lingua franca’ in scientific-academic contexts,” by identifying its deleterious effects on knowledge production, while arguing for “why we, as research communities in different fields and regions, should use multiple languages and varieties to promote transnational dialogue in scientific-academic contexts” (2022). Similarly, Suresh Canagarajah, at Penn State, expresses his hope this bibliodiversity amounts to a “gradual chipping away at power, to decolonize writing and scholarship as diverse communicative spaces, which will complement the parallel activism for large scale institutional and policy changes” (2022, p. 20). While the research hegemony of English is present among the JUOJS,

another linguistic feature of this collection is that a modest majority (54.5 percent) are publishing in more than one language (Figure 5).

Figure 5. Number of languages employed by journals using OJS (*N* = 22,382).



Note

1. Jupyter Notebook analysis and R visualization.

4.0 Discipline

4.1 Discipline Methodology

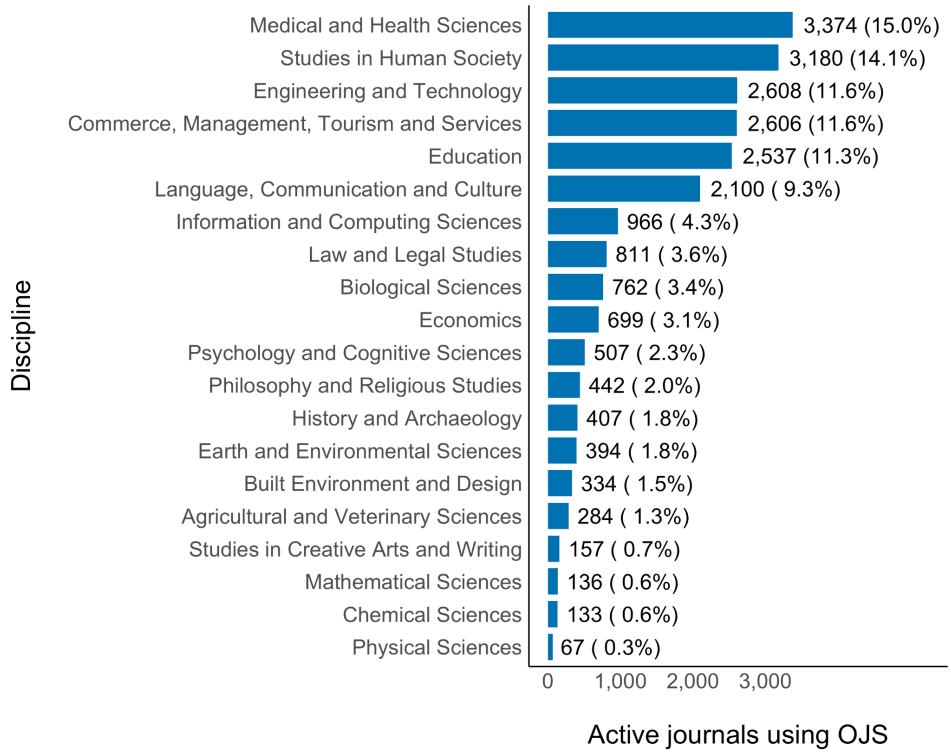
In an effort to see the diversity in research coverage found among JUOJS, we determined the different academic disciplines of each journal. To do so, we joined the title and abstract of the five most recent articles published in the subsample of 22,561 ISSN-verified journals to form a text input for each journal for a neural field of study classifier (Weber et al., 2020). Because this study classifier was trained on English-language data, all articles in journals that did not publish in English were first

translated into English using the Ubiquitous Knowledge Project's EasyNMT translator. The translated inputs were then passed into the field of study classifier, resulting in each journal being assigned the most probable field of study label (according to the ANZSRC system).

4.2 Discipline Results

The disciplinary distribution of journals reveals that social science research is the leading area among journals, although the STEM fields – led by the medical and health sciences, followed by engineering and technology – also play a major role by journal count (Figure 6). The STEM presence is notable where the expectation is that research conducted beyond the center represents a degree of cultural diversity that will occupy the main body of research activity (Figure 7). Where the social sciences and humanities are assumed to have a local significance that can be well served by researchers on site, the "universality" of the sciences frees them up for gravitating toward the center, with these results suggesting a more distributed basis for research in the sciences, technologies, engineering, and medicine (STEM).

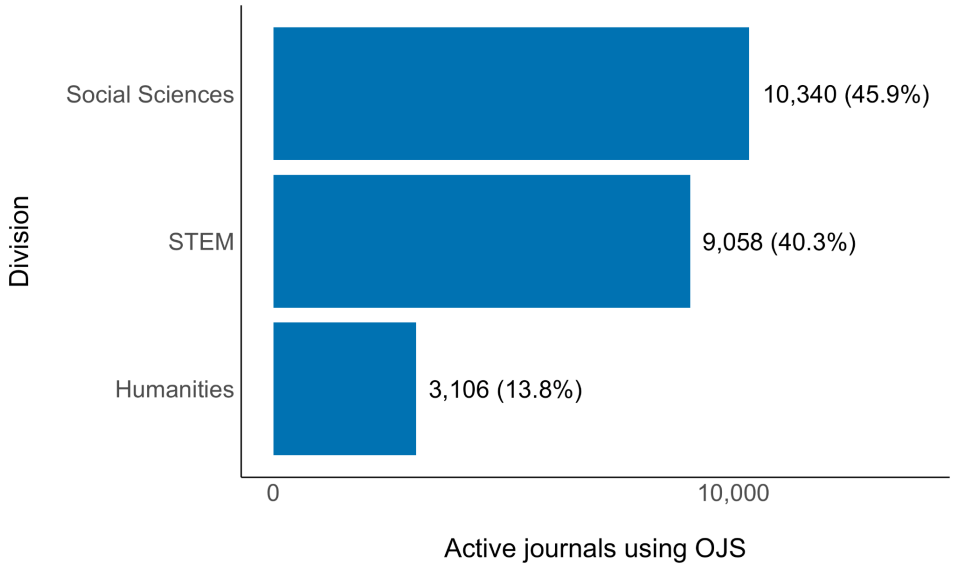
Figure 6. The disciplines of journals using OJS to publish in top 10 primary languages (*N* = 22,504).



Note

1. Jupyter Notebook analysis and R visualization.

Figure 7. The distribution of journals using OJS to publish in top-ten primary languages by their disciplinary coverage (*N* = 22,504).



Notes

1. Jupyter Notebook analysis and R visualization.

2. The top-ten languages, referring to the primary language of publication for journals, are identified in Table 2.

To compare the JUOJS disciplinary breakdown to the profile found in Scopus (with an overlap of less than four percent between them; see below), one finds a mixed picture. On the one hand, Scopus has a higher proportion of titles in (a) mathematics and (b) medicine and health sciences, while the set of JUOJS have a higher proportion in (a) education and (b) language, communication and culture (Table 2). Yet in other disciplinary areas the journal coverage of disciplines is roughly comparable between the JUOJS and Scopus two sets, such as (a) economics, (b) engineering and computer science, and (c) philosophy and religion (Table 2). One can imagine more advanced citational studies to determine the extent to which journals in these two sets draw on the other in their shared disciplinary interests.

Table 2. Disciplinary coverage by number of titles, comparing journals using OJS with journals in Scopus.

Field	Journals using OJS (*n* = 22,504 journals)	Scopus (*N* = 40,562 journals)
Economics	699 (3.1%)	1,335 (3.2%)
Education	2,537 (11.3%)	1,851 (4.2%)
Engineering and computer science	3,574 (15.9%)	6,290 (15.5%)
Language, communication & culture	2,100 (9.3%)	2,636 (6.0%)
Mathematics	136 (0.6%)	1,871 (4.6%)
Medicine and health sciences	3,374 (15.9%)	13,533 (33.4%)
Philosophy and religion	442 (2.0%)	1,153 (2.8%)

Note

1. Scopus was searched June 19, 2022, using its “subject area” filters.

5.0 Indexing and citation

5.1 Indexing and citation methodology

The visibility and circulation of the JUOJS were assessed by determining the presence of these journals in 11 research indexes, directories, and lists. The analysis of the Web

of Science Citation Reports and Cabells Predatory Reports involved drawing on Stanford University Library licensing agreements, with Cabells kindly determining the JUOJS in its database, based on its use of our data set. EBSCOHost, Scopus, OpenAlex, DOAJ, and Beall's List make their respective journal lists publicly available. Bianca Kramer ascertained the number of JUOJS in ROAD (2022), while SerpApi generously assisted in assessing Google Scholar's indexing of JUOJS. With each of the indexes, we then assess how well JUOJS are represented in the index by assessing overlap based on matched ISSNs, and/or matched journal URLs and domain names. The specific matching criteria used for each resource are also reported with Table 3.

5.2 Indexing and Citation Results

The marginalization of Global South research in the principal journal indexes, as part of the center-periphery colonial legacy, had been raised as early as 1998 by Ann María Cetto and Octavio Alonso-Gamboa, both at the National Autonomous University of Mexico (UNAM). That was the year they published a review of scholarly publishing in Latin America and the Caribbean that established how poorly the Web of Science covered the region, amounting to 0.5 percent of the journals indexed (Cetto & Alonso-Gamboa, 1998, p. 114). Cetto and Alonso-Gamboa were inspired to join forces with others across the region to form Latindex in 1997.

A quarter century after the Cetto and Alonso-Gamboa study, JUOJS represent only 1.1 percent of the journals in the Web of Science Citation Reports (Table 3).[14] The

[14] The Web of Science Citation Reports expanded its Core Collection in 2015 to include an Emerging Sources Citation Index, currently consisting of 7,800 titles, for which it will release its influential Impact Factors, beginning in 2023 (Cochran, 2022).

coverage of JUOJS in EBSCO and Scopus is only marginally better. That a majority of the JUJOS are included in Dimensions and OpenAlex is likely the result of an indexing model involving the automated aggregation of open resources, such as Microsoft Academic Graph (no longer maintained), Crossref, Unpaywall, DOAJ, ORCID, and PubMed. Google Scholar achieved the highest JUOJS' coverage, with 88.3 percent of the journals included, as a result of PKP working with this index since 2004 to improve the indexing of JUOJS.

Table 3. Journals by index, directory, and list, with proportion that use OJS (*N* = 25,671).

	Journals	JUOJS	Index %	OJS %
a) General research indexes				
Web of Science Citation Reports[a]	24,510	279	1.1%	1.2%
EBSCOHost[a]	17,874	771	4.3%	3.4%
Scopus[a]	41,957	1,646	3.9%	7.2%
Dimensions[a]	72,990	12,435	17.0%	54.5%
OpenAlex[a]	124,073	16,366	13.2%	63.8%
Google Scholar[c]	–	22,679	–	88.3%
b) Regional research index				
Latindex[a]	24,486	4,208	17.2%	66.6%[d]

c) Open access research indexes				
Directory of OA Journals (DOAJ)[a, b]	17,213	5,312	30.9%	20.7%
Directory of OA Resources (ROAD)[e]	37,333	10,976	29.4%	42.8%
d) “Predatory” journal lists				
Cabells Predatory Reports[a]	7,490	237	3.2%	1.0%
Beall's List[b]	38,849[f]	366	0.9%	1.4%

Notes

[a] Journal matches by ISSN for JUOJS (*N*=22,809) and those in the service’s set: Web of Science (mapped April 2022), EBSCOHost (mapped January 2022), Scopus (mapped January 2021), Dimensions (mapped January 2022), OpenAlex (mapped May 2022), Latindex (mapped June 2022), DOAJ (mapped June 2021), and Cabells Predatory Reports (mapped November 2021).

[b] Journal matches by URL for JUOJS (*N*=25,672) and those in the service’s set: Google Scholar (mapped June 2022 using domain name rather than journal URL), DOAJ (mapped March 2021), and Beall’s List (mapped January 2021).

[c] Google Scholar does not make its total journal count available or calculable.

[d] Latindex analysis includes only JUOJS published in this index’s participating countries (*N*=6,319).

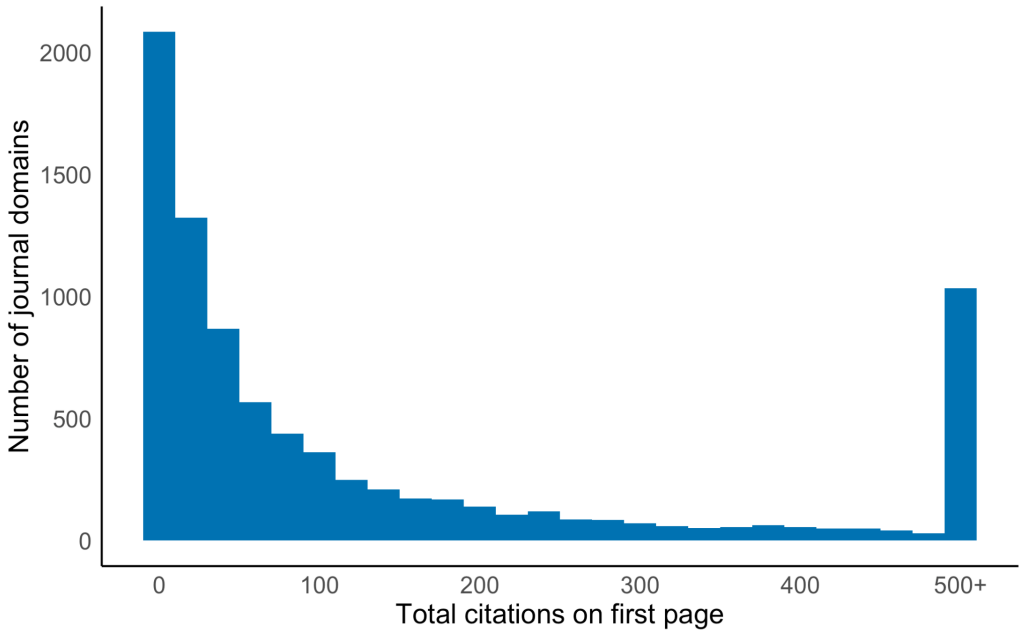
[e] Analysis performed by Bianca Kramer (2022).

[f] The journal total for Beall’s list of publishers and journals (*N*=39,849) was extrapolated from Chen’s (2019) sample.

Beyond indexing, we can also gain a sense of the circulation of the knowledge contained in JUOJS by examining the number of times they have been cited in the most comprehensive index for JUOJS – Google Scholar. We query Scholar using the domain

name extracted from each OJS journal URL, and log resultant research articles on the first page of results. We find that 22,679 JUOJS, hosted on 8,548 distinct domains, are present in Scholar. We also gain signal on the citational use for these journal domains. Out of the 8,548 JUOJS domains present in Scholar, 34 domains had over 10,000 citations, 565 journal domains had over 1,000 citations, and 1,015 journal domains had over 500 citations returned on the first page of their Scholar search results (see Figure 8 for complete distribution). Further, across all JUOJS domains present in Scholar, the mean of the total citations on the first Scholar page was 358.8 citations (median: 50.0). Still, 6.4 percent of the journal domains had no citations on that first page (with the most recent articles published by the JUOJS not less than 18 months earlier in 2020). While much remains to be done in establishing who is citing whom and where among JUOJS, this preliminary analysis provides a warrant for further investigating the nature and extent to which these journals are contributing to the research literature.

Figure 8. Citations for first ten articles listed for each domain hosting JUOJS in Google Scholar (*N* = 8,548 domains).



Notes

1. Rmarkdown Notebook Analysis and R visualization.

2. Google Scholar search conducted May 31, 2022.

5.2.1 Regional Index

Latindex is a regional index that draws on 17 national resource centers across Latin America, the Caribbean, and the Iberian Peninsula for its bibliographic data. It was started in 1997, as the Web of Science's underrepresentation of research from the region was becoming apparent (Cetto & Alonso-Gamboa, 1998). Of the 24,486 journals indexed in Latindex, 4,208 (17.2%) are JUOJS. On the other hand, fully two-thirds of the JUOJS published within this Iberoamerican region are indexed in Latindex.

5.2.2 Open Access Indexes

The Directory of Open Access Journals (DOAJ) is an independent organization supported by libraries, publishers, and others, that accepts open access journals for indexing based on their adherence to the criteria of scholarly standards. DOAJ is indexed for the 80 languages that the journals listed will accept articles in, although it is not clear how many of those languages in which an article has been published. The Directory of Open Access scholarly Resources is operated by the International Center for the registration of serial publications that manages ISSN (International Standard Serial Number) registration. It also has a series of scholarly standards for inclusion, while taking data from related sources, such as DOAJ, Latindex, and Scopus, which accounts for its larger size, with twice the number of JUOJS, while still representing a screening process.

5.2.3 “Predatory” Journal Indexes

The University of Colorado Denver librarian Jeffrey Beall’s designation of “potential, possible, or probable predatory scholarly open access publisher[s]” has played an outsize role in the discounting of Global South journals. Of the two prominent lists of so-called “predatory” journals, Beall maintained his blog post list from 2008 until 2017, at which time Cabells began to independently offer what is now Predatory Reports, which requires a subscription for access.[15] Beall’s 2017 list, which survives on the web, has 366 journals that are known to be using OJS, while Cabells’ current Predatory Report has 237, with the two lists sharing 82 titles in common. –

[15] In 2021, Stanford University Library paid $3,500/year for its Cabells Predatory Reports subscription at our request.

We have no doubts that some of these journals may well engage in deceptive practices. Yet the difficulty in determining that is a problem for scholarly publishing as it grows increasingly public and global. The designation *predatory* – which is intended to indict open access journals that charge fees while lacking scientific rigor – has itself been repeatedly criticized for its own want of rigor.[16] This has not prevented recent and repeated assertions that "predatory journals are a global threat" (Oviedo-García, 2021; Grudniewicz et al., 2019), with over 7,000 papers now addressing that threat, to judge from Google Scholar, reflecting a moral panic that cannot help but further undermine trust in this growing, larger world of science. A number of these papers have tended to conflate predatory with open access publishing (Krawczyk & Kulczycki, 2021). For those involved in developing scholarly publishing platforms, the question is how to provide readers with a better, more explicit (if still not perfect) means to assess the trustworthiness of all journals. We propose elsewhere raising the bar for journal's adherence to scholarly standards by having journals networked with industry organizations, such as ORCID and Crossref, to verify, for example, peer review (Khanna & Willinsky, in press).

6.0 Conclusion

This paper does not answer the question of how many journals there are in the world. It establishes, however, that the previously cited serval estimates of 30,000 fall woefully

[16] John Bohannon's notable experiment in determining deceptive journals, involving the submission of hoax study, found 18 percent of the titles selected from Beall's list redeem themselves by rejecting the submission, only to have it accepted by journals published by Elsevier, Wolters Kluwer, and SAGE (Bohannon, 2013).

short. Scholarly publishing is taking place on far more of a global scale than is commonly recognized in such estimates, as well as in studies that rely on, for example, the Web of Science Citation Index and Scopus. Dismissing that scale as a matter of too much research or out of an exaggerated concern over predatory publishing seems contrary to an interest in promoting the progress of science. For this reason, we call for a recalibration of scholarly communication that will help the world take advantage of the scale and growth of the research enterprise. This is about more than journal numbers. It concerns taking into account the significant bibliodiversity in scholarly communication reflected, in this case, by the 60 languages in which research is published from 136 countries. It involves the economics of scholarly communication. The study makes clear than an open source alternative to the commercial publishing services relied upon by profit and non-profit publishers has grown into what is arguably the most widely used journal platform in the world.[17] Such a platform is driving an OA diamond model of equitable access for authors, users, and readers across the full range of academic disciplines.[18]

[17] Reflecting the need take this economic alternative into account, Roger Schonfeld recently observed, under the theme of "keeping [scholarly publishing] infrastructure independent," that while "there are also open source alternatives, including Janeway and OJS," these alternatives "do not yet seem to have substantially impacted the market share of the commercial infrastructure providers" (Schonfeld, 2022).

[18] The OA diamond model may be the most effective challenge to "the more salient mechanisms through which the inequalities of knowledge production between the North and South are maintained," namely, the corporate sector's "market concentration, commodification, and monopolization" of journal publishing (Collyer, 2018, p. 60).

Still, this study represents a preliminary investigation in rescaling scholarly communication. It leaves many questions unanswered on how these publications operate within the larger world of research. How are the researchers publishing in JUOJS engaging and contributing to the work of others? There's a need to map the citation patterns to learn about how this work interacts within and beyond this set of journals. Further data analysis can establish whether this work is contributing to observed increases in the diffusion among works cited that is improving the efficiency with which research is circulating and being used.[19] Case studies may reveal whether it is having an influence on policy-making and professional practice. With the Public Knowledge Project and studies like this, we are aligned with the 2030 goal set by Research4Life, in conjunction with its 160 "publisher partners," which is to raise the "profile… and support for the Global South research publishing industry" (*Our vision*, 2022, p. 26).

Finally, we would note that the global scale of this publishing activity is in accord with an observation of Achille Mbembe at the University of the Witwatersrand – in paying tribute to Ngũgĩ wa Thiong'o's re-centering project of the last century – that "the decolonizing project is back on the agenda worldwide" (2019, p. 18, original emphasis). Now, whether this project is leading universities to become, as Mbembe envisions,

[19] "All these measures converge to demonstrate that citations are not becoming more concentrated but increasingly dispersed, and one can therefore argue that the scientific system is increasingly efficient at using published knowledge. Moreover, what our data shows is not a tendency towards an increasingly exclusive and elitist scientific system, but rather one that is increasingly democratic" (Larivière, Gingras, & Archambault, 2009, p. 861).

“platforms for the redistribution of different kinds of knowledges,” has yet to be established (p. 7, original emphasis). Yet what should be clear about this worldwide agenda, when it comes to the burgeoning research literature of the Global South, could be said to be summed up by Handel Kashope Wright and Yao Xiao at the University of British Columbia with their assertion that “in terms of ontology and epistemology and indeed the world order, decolonization helps us see that another world is possible” (2021). For our part, we hold that another world of scholarly communication – more broadly global, diverse, and inclusive – is already unfolding in ways that need not to be counted and studied for the benefit of all.

Acknowledgements

The authors wish to thank Anurag Acharya, Bianca Kramer, Ron Naoko, Kathy Kerns, SerpApi, and Cabells for their assistance in the preparation of this paper.

Competing interests

As acknowledged in the text, the authors are associated with the Public Knowledge Project, developer of Open Journal Systems, which is freely distributed as open source software (JB and SK serve as research associates, while JPA and JW are Co-Scientific Directors). In response, only is access is provided to both the data and the analysis code, as noted in Data and Code Availability, but subsequent data sets will continue to be made available as part of this open source software initiative.

Funding Information

This research was supported by the Graduate School of Education and the Khosla Family Professorship at Stanford University.

Data and Code Availability

The study's dataset has been made public (Khanna et al., 2021), and Jupyter Notebooks, linked to the corresponding figures below, provide the programs and scripts used to analyze the data (Ball, 2022; Khanna, 2022).

References

Action steps for rebuilding Ukraine's science, research, and innovation (2022), National Academies, Washington, https://www.nationalacademies.org/news/2022/06/action-steps-for-rebuilding-ukraines-science-research-and-innovation

Alatas, S. F. (2022). Political economies of knowledge production: On and around academic dependency. *Journal of Historical Sociology*, 35(1), 14-23. https://doi.org/10.1111/johs.12362

Alatas, S.H. (2000) Intellectual imperialism: definition, traits, and problems. *Southeast Asian Journal of Social Science*, 28(1), 23–45. https://doi.org/10.1163/030382400X00154

Alperin, J. P., K. Stranack, E. Hanson (2017). What is being published with OJS? and by whom? *PKP Scholarly Publishing Conference*, Montreal https://speakerdeck.com/jalperin/what-is-being-published-with-ojs-and-by-whom

Altbach, P. G. (1976). Publishing in Africa in the seventies: Proceedings of an International Conference on Publishing and Book Development, held at the University of Ife, Ile-Ife, Nigeria, 16-20 December 1973. *Library Quarterly: Information, Community, Policy*, 46(4), 459-461. https://www.jstor.org/stable/4306734

Altbach, P. G. (1981). The university as center and periphery. *Teachers College Record*, 82(4), 601-21. https://doi.org/10.1177/016146818108200412

Altbach, P. G. The role and nurturing of journals in the Third World. In P. G. Altbach & D. Tamtew (Eds.), *Knowledge dissemination in Africa: The role of scholarly journals* (1-12). Oxford: Bellagio Studies in Publishing, 1998.

Altbach, P. G. and de Wit, H. (2018, September 7). Too much academic research is being published, *University World News*. https://www.universityworldnews.com/post.php?story=20180905095203579

Amano T, Berdejo-Espinola V, Christie AP, Willott K, Akasaka M, Báldi A, et al. (2021) Tapping into non-English-language science for the conservation of global biodiversity. *PLOS Biology* 19(10): e3001296. https://doi.org/10.1371/journal.pbio.3001296

Ball, J. (2022), Open Journal Systems global. https://github.com/j-a-ball/ojs-global.

Beall's list of predatory journals and publishers, October 4, 2021. https://bealllist.weebly.com/

Berger, M. (2021). Bibliodiversity at the centre: Decolonizing Open Access. *Development and Change*, 52(2), 383-404. https://doi.org/10.1111/dech.12634

Bohannon, J. (2013, October 4). Who's afraid of peer review? *Science Magazine* 342(6154), 60-65. https://www.science.org/doi/10.1126/science.2013.342.6154.342_60

Bosman, Jeroen, Jan Erik Frantsvåg, Bianca Kramer, Pierre-Carl Langlais, Vanessa Proudman, *The OA diamond journals study: Exploring collaborative community-driven publishing models for open access* (March 2021) Sponsored by cOAlition S and Science Europe. https://doi.org/10.5281/zenodo.4558704

Canagarajah, S. (2022). Language diversity in academic writing: toward decolonizing scholarly publishing. *Journal of Multicultural Discourses*, 1-22. https://doi-org.stanford.idm.oclc.org/10.1080/17447143.2022.2063873

Cetto, A. M., & Alonso-Gamboa, O. (1998). Scientific and scholarly journals in Latin America and the Caribbean. In P. G. Altbach & D. Teferra (Eds.), *Knowledge dissemination in Africa: The role of the scholarly journal*, Bellagio Studies in Publishing, 99-126.

Chan, L. (2018, December 12). Open Access, the Global South and the politics of knowledge production and circulation, *Open Insights* [blog]. https://www.openlibhums.org/news/314/

Cloudscene. (2022, January 26). Number of data centers worldwide in 2022, by country [Graph]. In Statista. Retrieved July 15, 2022. https://www.statista.com/statistics/1228433/data-centers-worldwide-by-country/

Cochran, A. (2022). The end of Journal Impact Factor purgatory (and numbers to the thousandths). *Scholarly Kitchen*, https://scholarlykitchen.sspnet.org/2022/07/26/the-end-of-journal-impact-factor-purgatory-and-numbers-to-the-thousandths/

Collyer, F. M. (2018). Global patterns in the publishing of academic knowledge: Global North, global South. *Current Sociology*, 66(1), 56-73.https://doi.org/10.1177/0011392116680020

Collyer, F., Connell, R., Maia, J., & Morell, R. (2019). *Knowledge and global power : Making new sciences in the South*. Monash University Publishing.

Dados, N., & Connell, R. (2012). The global south. *Contexts*, 11(1), 12-13. https://doi-org.stanford.idm.oclc.org/10.1177/1536504212436479

Edgar, B. D., & Willinsky, J. (2010). A survey of scholarly journals using Open Journal Systems. *Scholarly and research communication*, 1(2), https://www.nature.com/articles/d41586-021-00899-y

Elnathan, R. (2021, April 1). English is the language of science — but precision is tough as a non-native speaker, *Nature*, 2021, https://www.nature.com/articles/d41586-021-00899-y

Gottschalk, C. M., & Desmond, W. F. (1963). Worldwide census of scientific and technical serials. *American Documentation*, 14(3), 188-194, https://doi.org/10.1002/asi.5090140303

Grudniewicz, A., D. Moher, […] Manoj M. Lalu. (2019, December 11). Predatory journals: no definition, no defence, *Nature*, https://www.nature.com/articles/d41586-019-03759-y

Haug, Sebastian. (2021, September 28). What or where is the “Global South”? A social science perspective. *LSE Impact Blog*, https://blogs.lse.ac.uk/impactofsocialsciences/2021/09/28/what-or-where-is-the-global-south-a-social-science-perspective/

Hawkins, D. (2022). Don’s conference Notes. Charleston in between 2022, Charleston Hub. https://www.charleston-hub.com/2022/05/dons-conference-notes-charleston-in-between-2022/

Hunter-Konos, A. (2015, August 21). Hidden bias of science’s universal language, *Atlantic,*

https://www.theatlantic.com/science/archive/2015/08/english-universal-language-science-research/400919/

Khanna, S. (2022). Assessing OJS overlaps with scientometric databases. https://github.com/j-a-ball/ojs-global

Khanna, S. and J. Willinsky (in press). What those responsible for open infrastructure in scholarly communication can do about possible predatory practices. In I. Fazel & P. Habibie (Eds). *Predatory practices in scholarly communication and publishing: Causes, forms, implications, and solutions*. Routledge, preprint https://preprints.scielo.org/index.php/scielo/preprint/view/3474

Kramer, B. (2022). *PKP dataset 202111 matched to ROAD, DOAJ*. Unpublished analysis, https://docs.google.com/spreadsheets/d/17UihGAMruHVecbzwc-VA6QLriCaPD8-7kzWOzsmDbxM/edit?usp=sharing

Krawczyk, F., & Kulczycki, E. (2021). How is open access accused of being predatory? The impact of Beall's lists of predatory journals on academic publishing. *Journal of Academic Librarianship*, 47(2), 102271. https://doi.org/10.1016/j.acalib.2020.102271

Larivière, V, Yves Gingras, and Éric Archambault. The decline in the concentration of citations, 1900–2007, *Journal of the American Society for Information Science and Technology* 60, no. 4 (2009): 858-862. https://doi-org.stanford.idm.oclc.org/10.1002/asi.21011

Mbembe, Achille. (2015) *Decolonizing knowledge and the question of the archive*. Public lectures given at the Wits Institute for Social and Economic Research (WISER),

https://wiser.wits.ac.za/system/files/Achille%20Mbembe%20-%20Decolonizing%20Knowledge%20and%20the%20Question%20of%20the%20Archive.pdf

Muddit, A. (2020, October 28). In search of equity and justice: Reimagining scholarly communication, *Scholarly Kitchen* [blog] https://scholarlykitchen.sspnet.org/2020/10/28/in-search-of-equity-and-justice-reimagining-scholarly-communication/

Navarro, F., Lillis, T., Donahue, T., Curry, M. J., Ávila Reyes, N., Gustafsson, M., Zavala, V., Lauría, D., Lukin, A., McKinney, C., Feng, H., & Motta-Roth, D. (2022). Rethinking English as a lingua franca in scientific-academic contexts. A position statement. *Journal of English for Research Publication Purposes*, 3(1), 143-153. https://benjamins.com/catalog/jerpp.21012.nav

Oviedo-García, M. A. (2021). Journal citation reports and the definition of a predatory journal: The case of the Multidisciplinary Digital Publishing Institute (MDPI). *Research Evaluation*, 30(3), 405-419. https://doi.org/10.1093/reseval/rvab0201

Ondari-Okemwa, E. (2007). Scholarly publishing in sub-Saharan Africa in the twenty-first century: Challenges and opportunities. *First Monday*, 12(10). https://doi.org/10.5210/fm.v12i10.1966

Our vision for 2030 (2022). Research4Life strategic plan, https://www.research4life.org/news/vision-to-2030-research4life-strategic-plan/

Pacchioni, Gianfranco (2018). *The overproduction of truth: Passion, competition, and integrity in modern science.* Oxford University Press.

Raoni, J. (2021). MARC Countries <=> ISO. https://github.com/jonasraoni/marc-to-iso

Schonfeld, R. (2022, August 13). Keeping publishing infrastructure Independent. *Scholarly Kitchen* https://scholarlykitchen.sspnet.org/2022/08/15/keeping-publishing-infrastructure-independent/

STM Report: STM Global Brief 2021 – Economics and market size (2021). International Association of Scientific, Technical and Medical Publishers, https://www.stm-assoc.org/2021_10_19_STM_Global_Brief_2021_Economics_and_Market_Size.pdf

Thiong'o, Ngũgĩ wa. (1993). *Moving the center: The struggle for cultural freedoms*. James Currey, London.

Van Noorden, R. (2019, May 15). Indonesia tops open-access publishing charts. *Nature*. https://www.nature.com/articles/d41586-017-08838-6

Vera-Baceta, M. A., Thelwall, M., & Kousha, K. (2019). Web of Science and Scopus language coverage. *Scientometrics*, 121(3), 1803-1813. https://doi.org/10.1007/s11192-019-03264-z

Weber, T., Kranzlmüller J., Fromm, M., and Tavares de Sousa, N. (2020). Using supervised learning to classify metadata of research data by field of study. *Quantitative Science Studies* 1 (2): 525–550. https://doi.org/10.1162/qss_a_00049. See also Field of Study Classification (FOSC). https://github.com/tgweber/fosc/

Wickham, H. (2016). Data analysis. In *ggplot2* (pp. 189-201). Springer, Cham.

Willinsky, J. (2017) . *The intellectual properties of learning: A prehistory from Saint Jerome to John Locke*. Chicago University Press. https://intellectualproperties.stanford.edu/download

Wright, H. K., & Xiao, Y. (2021). Decolonisation and higher education: Theory, politics and global praxis. *Postcolonial Directions in Education* 10(1): 23-50. https://www.um.edu.mt/library/oar/bitstream/123456789/78311/1/PDE10_1_A2.pdfb

This preprint was submitted under the following conditions:

- The authors declare that they are aware that they are solely responsible for the content of the preprint and that the deposit in SciELO Preprints does not mean any commitment on the part of SciELO, except its preservation and dissemination.
- The authors declare that the necessary Terms of Free and Informed Consent of participants or patients in the research were obtained and are described in the manuscript, when applicable.
- The authors declare that the preparation of the manuscript followed the ethical norms of scientific communication.
- The authors declare that the data, applications, and other content underlying the manuscript are referenced.
- The deposited manuscript is in PDF format.
- The authors declare that the research that originated the manuscript followed good ethical practices and that the necessary approvals from research ethics committees, when applicable, are described in the manuscript.
- The authors declare that once a manuscript is posted on the SciELO Preprints server, it can only be taken down on request to the SciELO Preprints server Editorial Secretariat, who will post a retraction notice in its place.
- The authors agree that the approved manuscript will be made available under a Creative Commons CC-BY license.
- The submitting author declares that the contributions of all authors and conflict of interest statement are included explicitly and in specific sections of the manuscript.
- The authors declare that the manuscript was not deposited and/or previously made available on another preprint server or published by a journal.
- If the manuscript is being reviewed or being prepared for publishing but not yet published by a journal, the authors declare that they have received authorization from the journal to make this deposit.
- The submitting author declares that all authors of the manuscript agree with the submission to SciELO Preprints.