DADOS
dados.iesp.uerj.br

[Artigos Originais]

OLIVA: The Scientific output in journals edited in Latin America. Disciplinary Diversity, Institutional Collaboration, and Multilingualism in SciELO and Redalyc (1995-2018)*

Fernanda Beigel[1]

[1]Researcher at INCIHUSA-CONICET and Professor of Sociology at the Universidad Nacional de Cuyo. Director of the Centro de Estudios de la Circulación del Conocimiento (FCPyS-UNCuyo). Mendoza, Argentina.
Email: fernandabeigel@gmail.com ORCID: https://orcid.org/0000-0002-7996-9660

Abel L. Packer[2]

[2]Director of the SciELO / FAPESP (Scientific Electronic Library Online) Program. São Paulo, SP.
Email: abel.packer@scielo.org ORCID: https://orcid.org/0000-0001-9610-5728

Osvaldo Gallardo[3]

[3]Postdoctoral Fellow at IMESC-CONICET. Professor of History at the Universidad Nacional de Cuyo. Researcher of the Centro de Estudios de la Circulación del Conocimiento (FCPyS-UNCuyo). Mendoza, Argentina.
Email: osvaldogallardo87@gmail.com ORCID: https://orcid.org/0000-0003-0662-2196

Maximiliano Salatino[4]

[4]Researcher at INCIHUSA-CONICET and Professor of Political Science at the Universidad Nacional de Cuyo. Researcher of the Centro de Estudios de la Circulación del Conocimiento (FCPyS-UNCuyo). Mendoza, Argentina.
Email: maxisalatino@gmail.com ORCID: https://orcid.org/0000-0001-5573-9599

DOI: https://doi.org/10.1590/dados.2024.67.1.307x

Dataset: https://ri.conicet.gov.ar/handle/11336/175850

*We are deeply grateful for the essential comments by this article's two peer reviewers, and for Dominique Babini's thoughtful review. This research was possible thanks to the financial support received from the Universidad Nacional de Cuyo (PROYECTO SIIP TIPO 4-038) and the Agencia I+D+I-Ministerio de Ciencia, Tecnología e Innovación, Argentina (PICT 2017-2647).

Resumo

OLIVA: A Produção Científica Indexada na América Latina. Diversidade Disciplinar, Colaboração Institucional e Multilinguismo em Scielo E Redalyc (1995-2018)

Este artigo apresenta os resultados do *Observatorio Latinoamericano de Indicadores de eVAluación* (OLIVA) que visa visibilizar e valorizar a produção científica indexada na América Latina e Caribe. O trabalho aborda a produção em acesso aberto das revistas indexadas no SciELO e Redalyc, com base na consolidação de uma única base de dados, com o número total de 1.720 revistas (de 15 países), 908.982 documentos e 2.591.704 autores/as. Também analisa a diversidade disciplinar e as tendências da colaboração nacional e internacional. Por fim, apenas para o caso do Brasil e da SciELO, analisa-se a colaboração intranacional. O estudo conclui que há um predomínio das revistas diamante, das como instituições editoras universitárias e de formas de circulação a multiescalar. Estas caraterísticas, mesmo assim a diversidade linguística e disciplinar, podem contribuir de forma muito eficaz para as necessidades atuais de comunicação científica em tempos de ciência aberta.

Palavras-chave: revistas indexadas; SciELO; Redalyc; colaboração nacional; colaboração internacional

Abstract

OLIVA: The Indexed Scientific Output in Latin America. Disciplinary Diversity, Institutional Collaboration, and Multilingualism in SciELO and Redalyc (1995-2018)

This article presents the results of the Latin American Observatory of eVAluation Indicators (OLIVA, its Spanish acronym) which aims to contribute to the visibility of indexed scientific output in Latin America and the Caribbean and enhance its value in evaluation systems. This study addresses the production published in open access by journals indexed in SciELO and Redalyc, based on a single database of a total of 1,720 journals (from 15 countries), 908,982 documents and 2,591,704 authors. It also highlights its disciplinary diversity, and trends in national and international research collaboration. Finally, only for the case of Brazil and SciELO, intranational collaboration is analyzed. The study concludes that there is a predominance of diamond journals, of university publishing institutions and of multiscalar forms of circulation. These characteristics, even with linguistic and disciplinary diversity, can contribute very effectively to the current needs of science communication in times of open science.

Keywords: indexed journals; SciELO; Redalyc; national collaboration; international collaboration

Résumé

Oliva : Production Scientifique Indexée en Amérique Latine. Diversité Disciplinaire, Collaboration Institutionnelle et Multilinguisme Chez Scielo et Redalyc (1995-2018)

L'article présente les résultats de l'Observatoire Latino-Américain des Indicateurs d'Évaluation (OLIVA), qui visent à contribuer à rendre visible la production scientifique indexée en Amérique Latine et dans les Caraïbes, afin de promouvoir son appréciation dans les systèmes d'évaluation. Cet article aborde la production publiée en libre accès par des revues indexées dans SciELO et Redalyc, à partir de la consolidation d'une base de données, sans recouvrement, avec un total de 908.982 documents, 1720 revues (15 pays) et 2.591.704 auteurs. La diversité disciplinaire de cette production est analysée, ainsi que les tendances en matière de collaboration nationale, régionale et internationale. Enfin, la collaboration intra-nationale a été analysée uniquement dans le cas du Brésil et de SciELO. L'étude conclut qu'il y a une prédominance des revues « diamant », des institutions d'édition universitaire et des formes de circulation multiscalaires. Ces caractéristiques, même avec la diversité linguistique et disciplinaire, peuvent contribuer très efficacement aux besoins actuels de la communication scientifique à l'heure de la science ouverte.

Mots-Clés: revues indexées; SciELO; Redalyc; collaboration nationale; collaboration internationale

Resumen

Oliva: La Producción Científica Indexada en América Latina. Diversidad Disciplinar, Colaboración Institucional y Multilingüismo en Scielo y Redalyc (1995-2018)

Este artículo presenta los resultados del Observatorio Latinoamericano de Indicadores de eVAluación (OLIVA), que buscar visibilizar y promover la valoración de la producción científica indexada en América Latina y el Caribe. Se aborda la producción publicada en acceso abierto e indexada en SciELO y Redalyc, a partir de la construcción de una base de datos consolidada y que incluye información de 1.720 revistas (de 15 países), 908.982 documentos y 2.591.704 autores/as. Se analiza la diversidad disciplinar de esta producción, así como las tendencias de la colaboración nacional, regional e internacional. También, sólo para el caso de Brasil y de SciELO, se analiza la colaboración intra-nacional. Los resultados arrojan un predominio de revistas diamante, de instituciones editoras universitarias y de formas de circulación multiescalar. Estos rasgos, así como la diversidad lingüística y disciplinar, pueden contribuir de modo eficaz a las necesidades de la comunicación científica en tiempos de ciencia abierta.

Palabras-clave: revistas indexadas; SciELO; Redalyc; colaboración nacional; colaboración internacional

Introduction

The Latin American academic circuit is a constellation of national academic communities and regional networks formed in the mid-20th century with the support of intergovernmental organizations, cooperation agencies, and foundations that promoted scientific information management as a key element for development. This led to the emergence of regional institutions collaborating closely in cataloging and constructing bibliographic indexes to disseminate locally produced scientific knowledge (Rodríguez Garcia, 2020). By 1967, building information centers such as BIREME and research networks such as CLASCO constituted a significant step forward. They soon became major regional repositories for disseminating Latin American scientific output (Packer, 2005; Vessuri, 1994). Efforts to professionalize scientific publishing were driven by the creation of indexing systems led by regional centers affiliated with United Nations agencies (IMLA-LILACS, DOCPAL, REPIDISCA, AGRINTER-SIDALC) and by large public universities, most notably the Universidad Nacional Autónoma de México (responsible for the first indexes Clase, Periódica, and afterward Latindex and BIBLAT). Universities, national and inter-governmental organizations created and operated all these systems to improve the bibliographic exchange of the region's scientific and technical output. With the pioneering emergence of the regional open access journal platforms, SciELO in 1998 and Redalyc in 2005, an infrastructure supporting digitized journals and indexing web services was born. Furthermore, these services created a journal quality certification system with Latindex that focused on peer review and best practices in editing. This system resulted in a progressive professionalization of the journals, thus establishing an editorial hallmark in the region characterized by academic quality and non-commercial open access. With a strong public character and the commitment of many governments, these portals and indexing services today represent a fundamental space for open science to develop (Vessuri, Cetto, Guédon, 2014; Beigel, 2019; Packer, 2020; Banzato, Salatino, 2020).

The global report on diamond journals (Bosman et al., 2021) recently published by OPERAS shows the importance of journals that do not charge for publishing or reading and points to the role of Latin America in the edition of 25% of these total publications worldwide. Indeed, open access journals are mostly published by universities and full-time professors, assisted by technical support teams centralized in libraries, manage them. An important aspect in consolidating university publishing in the region

is the massive adoption of the open source Open Journal System (OJS-PKP) to manage editing, peer review and publication of journals.[1] In addition to university journals, there are also those of scientific societies and professional associations that the university infrastructure and personnel help operate: with digitization, by creating permalinks, acquiring DOI's, and doing the XML markup of texts. Although more support is needed, this core institutional and governmental support explains the existence and growth of Latin American and Caribbean journals.

The vibrancy of this regional circuit is now more globally visible given that the flow of scholarly communication has been improved by several mega indexes such as Google Scholar, Dimensions, Lens, and the federation of repositories LA Referencia, which covers a greater share of Latin American and Caribbean output. However, these journals still have no significant impact on global S&T reports or on evaluations of scientific careers, projects, and institutions in which the Impact Factor and other indicators from traditional databases, such as the Web of Science and Scopus, continue to prevail. This harms the value of the journals indexed in the region as it drives many research groups to choose journals managed by the dominant publishers in the industry. Historically, this choice by the academic community has been shaped by the deeply rooted belief in the significance/transparency of these indicators (Martinovich, 2020). It also highlights the gap between the ability of the region to produce and publish quality science and the journals' ability to gain international legitimacy (Packer and Meneghini, 2007). Several studies have already pointed out the limitations of these traditional indicators, which several structural conditions favor and that generated a series of abuses and distortions. These changes affected researchers' creativity and even concern for the social significance of science (Guédon, 2011; Gingras, 2016; Ráfols, 2019). Thus, policies of internationalization justified by the need to climb in the rankings result in many institutions having restricted autonomy and a weakened capacity to engage with the research environment.

The development of a global consensus on the need for open access to scientific literature has brought increasing attention to Latin American journals and indexing systems. Basson et al. (2021) analyze the proportion of open access articles available in Dimensions, which had 46.6% of its documents in open access by 2021. They note that this percentage rises to 63.5% for documents published by Latin American and Caribbean journals. There are few studies that, in terms of documents, explore the breadth and diversity of published output in the region, apart from the

aforementioned mainstream databases. An important precedent is the study by Miguel (2011) that compared the 2005-2009 coverage of Latin American and Caribbean journals within SciELO, Redalyc, and SCOPUS. The estimated volume of scientific output within these three sources revealed differences with respect to the coverage in the Latindex catalog[2]. The distribution of disciplines within SCOPUS and SciELO were more balanced, while Redalyc showed greater inclination toward the Social Sciences and Humanities.

Vuotto, Di Césare and Pallotta (2020) analyzed 17 bibliographic databases and verified that, today, there are still few that openly offer their elements in a common format and structure to allow output at the document level to be analyzed. For this reason, the available empirical studies up to now have only been able to analyze journal collections separately offered by each indexing database. Scopus has 890 journals published in the region, and Web of Science has 223. As of June 2021, SciELO had 1,358 active indexed journals also replicated in the SciELO Citation Index in the Web of Science platform. This allows for the evaluation of articles' performance by way of citations. Redalyc has its own information system and indicators for its collection of 1,415 journals. These collections share many journals, as we will see shortly. There is also Biblat, with its own cataloging system and a large number of complete records at the document level. However, it is not currently possible to work with combined data from the region's three indexing databases because they are not interoperable. We must add LA Referencia to these sources, which is a significant database that gathers the scientific output of 790 institutions from 12 Ibero-American countries. Currently, it has more than three million documents. However, it does not have advanced searching by journal and, because it originates from the repositories of individual institutions, there are overlaps of data with former indexing services.

Furthermore, there are thousands of journals that do not have metadata at the document level. This affects the possibility of analyzing the output published in Latindex journals, which continues to be – as Miguel already observed in 2011 – the most comprehensive cataloging database. In addition, institutional affiliations are not standardized, nor are the disciplinary classification of journals; nor is the use of unique identifiers of authors and institutions such as ORCID and ROR sufficiently widespread among the databases. All these further limit studies at the document level. Consequently, one can only observe this large body of documents published within the region from a limited and fragmented approach.

It is in this context that the OLIVA project (Latin American Observatory of eVAluation Indicators)[3] emerged as an initiative of the *Centro de Estudios de la Circulación del Conocimiento* (CECIC-*Facultad de Ciencias Políticas y Sociales*). The project was approved and funded in early 2019 by the Universidad Nacional de Cuyo. The main purpose of this project is to shed light on the abundant scientific output published in the region, and to develop indicators to show how valuable these journals are to assess individual careers and academic institutions. For the first phase of the project, the SciELO/FAPESP Program of Brazil, Redalyc (at the Universidad Autónoma del Estado de México -UAEM Mexico), and the Consejo Latinoamericano de Ciencias Sociales (CLACSO) were invited to collaborate. This paper presents the results of this first phase of the project. The second phase foresees the incorporation of documents from journals indexed in Latindex Catalog 2.0 and Biblat. A pilot is already underway within these institutions.

Within OLIVA, a database of 1,720 scientific journals published in Latin America and the Caribbean and indexed in SciELO and/or Redalyc was built. The study was organized on two levels: at the journal level and at the document level. The objective was to map out this circuit of quality scientific literature, addressing the existing forms of co-authorship, disciplinary characteristics, and languages of publication. In the first part of this article, we delve into describing journals in the database constructed at this stage (SciELO and Redalyc). We then move onto the journals' publishing institutions and highlight their predominantly public nature and self-management by the academic community. The journals' disciplines were empirically classified based on the information found on their respective websites. This was done due to the existing limitations in classifying documents based on the disciplinary catalog offered by the indexing databases. It revealed that the scientific output is broad in scope, with half of the indexed documents being from the Social Sciences and Humanities, and the other half from Engineering, Natural Sciences, and Health and Biological Sciences. We then analyzed the patterns of co-authorship and the different levels of collaboration – global, regional (LA) and national. We also analyzed the language patterns of the corpus, noting the growth of journals publishing in English and a multilingualism emerging in journals starting to publish in several languages.

Finally, we focus on the case of Brazil. Given its significant weight, with 50% of articles published in the database and a high tendency toward co-authorship among its authors, we contrast this with a more detailed

analysis of the SciELO Brazil collection (only journals from Brazil). There, we find that there is a strong tendency of scientific collaboration between authors affiliated with the same country but belonging to universities from different states. Thus, a picture emerges that questions the preconceived notion that collaboration between authors of the same country signifies academic inbreeding and reveals the complexity of the national academic space and its current dynamics.

Sources and methods

The OLIVA project database is made up of the collected data of journals indexed by SciELO and Redalyc, and revised data from the published documents. Each of these indexing systems collaborated by providing its historical database up to June 2019. The data were subjected to data cleaning and deduplication of common journals. The result was an integrated bibliographic database for statistical purposes and is composed of the following elements at the document level: type of document, date of publication, language, and author's country. It does not include data on citations that do not fall within the project's scope.

The construction of this integrated database faced significant limitations. As already been pointed out, the main limitation was the lack of a common identification system among indexing platforms for authors, journals, and institutional affiliations. There was no information on gender, as is usually the case with scientific documents and bibliographic records. An important difficulty was compiling the various indexed, published documents in journals pertaining to both indexing services.

In view of all this, we opted for a process of journal revision through a manual and meticulous procedure. First, we identified the journals present only in SciELO or Redalyc, and aggregated the metadata of these documents – articles, reviews, and editorials – both from active journals at the time of the study as well as those indexed at an earlier time. Journals common to both indexing platforms could have different gaps or periods of coverage in each of the platforms. Thus, we compared the consistency and completeness of data and, for each case, we preserved those that had the widest temporal range and the largest number of documents. To clarify, we only worked with journals published in Latin America and the Caribbean in this phase of the project.[4]

Regarding the disciplinary classification of journals, we faced the challenge of indexing databases using different classifications, even for grouping scientific areas. Thus, the same journal could be classified in different disciplines and consequently in different areas. The OLIVA database used a classification based on the OECD and Frascati Manual criteria. One or more disciplines was assigned to each journal through an empirical survey based on the description in the "Scope" or "About" section of a journal's official website. The journals were then grouped into eight subject areas: 1) Agricultural Sciences; 2) Social Sciences; 3) Humanities; 4) Engineering and Technology; 5) Medical and Health; 6) Natural and Exact Sciences; 7) Multidisciplinary; and 8) Multidisciplinary-Social Sciences & Humanities (SSH). The existence of two areas called "Multidisciplinary" is because there are a significant number of journals in the OLIVA database that are diversified and comprehensive in scope, crossing broader disciplinary boundaries. Thus, we included journals that belong to at least two subject areas, combining "hard" sciences with SSH in group 7. In group 8, we included journals that combine subject areas only within the Social Sciences and Humanities. This primary classification allowed us to reduce inconsistencies and to assess the disciplinary distribution of analyzed journals more accurately.

Information on the publishers of each journal as well as the APC information also came from each journal's website.

The scope of the Oliva *corpus* and the institutional attachment of publishers

the OLIVA database, constructed from the collections of Redalyc and SciELO, contains 908,982 records of documents from 1,720 journals published between 1909 and May 2019 (see Table 1). The main type of document is the article, which composes 87% of records. However, this percentage varies among subject areas. There is an average of 3.1 authors per document and 3.3 per article. Given that multiple authorship is more commonplace in the Exact and Natural Sciences, Engineering and Biological Sciences, data already indicate that we are dealing with an output that is broad in disciplinary terms. One might think that journals in the region publish only in the vernacular languages and therefore would be mostly from SSH, since they are accustomed to publishing in local languages. However, we will later see that this is a multilingual corpus.

Journals edited in Brazil represent 29.4% of the total and account for 48.9% of the total number of documents and 50.1% of the articles. This predominance of Brazilian output even more prominent in the number of authors, given the growing trend in the country of collaboration between 3 or more researchers, including in the social sciences and humanities. In terms of number of articles, journals from Mexico, Colombia, Chile, Cuba and Argentina follow Brazil. The journals from the rest of the region compose only 8.7% of articles.

Table 1

Basic data on the OLIVA *corpus*. Number of journals, documents, articles, and author records by country.

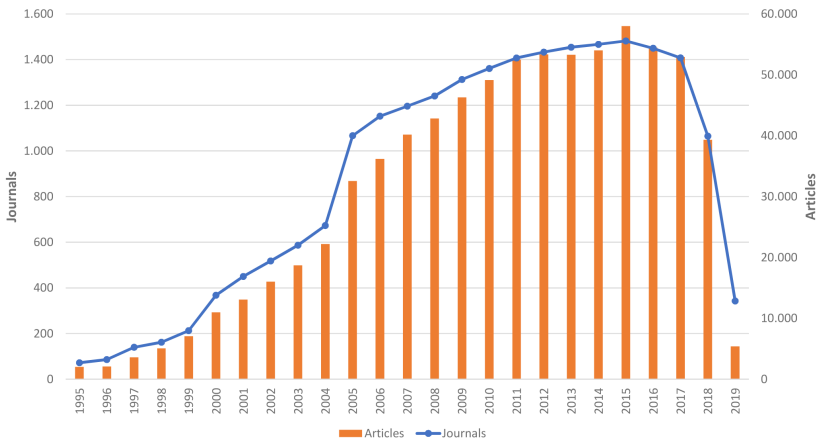
Country of publication	Journals	Documents	Document authors	Articles	Article authors
Brazil	506	444,332	1,579,011	396,293	1,476,492
Colombia	291	102,762	241,335	90,530	224,630
Mexico	283	120,475	299,471	100,355	273,002
Argentina	167	46,237	121,052	35,919	104,093
Chile	144	69,095	199,308	57,032	178,521
Venezuela	97	34,939	92,097	30,161	85,814
Cuba	82	46,052	160,371	41,621	149,736
Costa Rica	48	16,313	35,322	14,816	33,199
Peru	37	13,902	40,558	11,773	37,066
Uruguay	25	4,756	14,914	3,680	12,866
Bolivia	22	4,491	10,080	3,629	8,767
Ecuador	11	3,440	5,631	2,877	5,002
Puerto Rico	5	1,253	1,841	816	1,382
Panama	1	408	427	333	341
Dominican Republic	1	527	877	469	793
Total	1,720	908,982	2,802,295	790,304	2,591,704

Source: The OLIVA database as of June 2019. Note: the column "Journals" includes active and inactive journals indexed by the time the documents were published; the column "Documents" includes articles and other types of literature; the columns "authors" refer to the number of occurrences of authors in the journals published in each country, not the number of different authors in that country.

Concerning the temporal range of the available data, the corpus reveals more than a century of Latin American scientific outputpublished in indexed journals The oldest document dates back to 1909 and the most recent one to May 2019. However, 84.2% of the articles fall within the period 2005- 2018. There are many fewer documents in 2019 (0.67% of articles) as data collection took place between May and June of that year, which represents only a share of the year. The earliest documents are from the *Memórias do Instituto Oswaldo Cruz,* which has 582 records between 1909 and 1939. Then, we found the first documents of the *Revista Chilena de Pediatría* appearing in 1940, *Bragantia* in 1941, *Arquivos de Neuro-Psiquiatria* in 1943 and *Anais da Escola Superior de Agricultura Luiz de Queiroz,* today's *Scientia Agricola,* in 1944. The first appearances of journal documents from the remaining countries are in 1969 (Colombia), 1974 (Mexico), 1993 (Argentina and Cuba), 1995 (Costa Rica and Peru), 1984 – *Revista de la Facultad de Medicina* (Venezuela), 2000 (Ecuador, Puerto Rico, Dominican Republic and Uruguay), 2001 (Bolivia) and 2005 (Panama). The first journal in the Social Sciences and Humanities recorded in the database is the *Revista de Administração de Empresas* (1961) followed by Colombia's *Revista Latinoamericana de Psicología* (1969).

There are fewer than 100 documents per year until 1942; fewer than 1,000 until 1990; fewer than 10,000 until 1999. The highest number is in 2015, with 58,025 articles. These values should be interpreted with caution as they do not necessarily represent the history of Latin American and the Caribbean journals. Indeed, we were able to verify that in most of the older journals, the digitization and availability of the volumes prior to 1998 is patchy. There is only information for 78 journals and it corresponds to documents published before 1998 – the year SciELO was launched – so that the inclusion of documents prior to a journal's indexing year seems exceptional.

Figure 1
The OLIVA *Corpus* (1995-2019): journals with at least one article published per year; total articles published per year



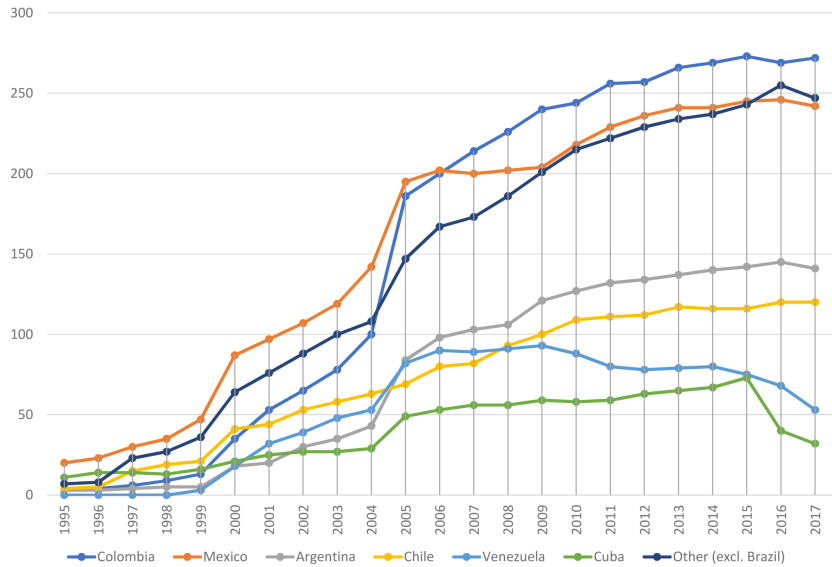
Source: The OLIVA database. The drop in the number of records in recent years is due to the data harvesting carried out in the first half of 2019.

The most notable growth in the number of journals and articles occurred with the launching of Redalyc in 2005, which began to index journals especially from the Social Sciences. As seen in Figure 1, the peak in 2015 shows a sharp increase relative to some flattening of the article curve beginning in 2011. The number of articles and journals per year decreases in 2018 because there is a delay between publishing and incorporating new issues into the Redalyc and SciELO databases. This delay varies widely from one publisher to another and among countries. It is due to a multitude of reasons, such as delays in publication or in managing databases, difficulties in XML-JATS markup, or how publishing institutions prioritize including new issues in each database.

Accordingly, the evolution of the journals shows two significant increases: one between 1998 and 2000 and another in 2005, the years when SciELO and Redalyc were launched. Journals published in Brazil make a significant contribution because there is a jump in the number of journals precisely in 1998. The curve continues its upward trend, reaching 198 journals in 2004 and 324 in 2005. The curves of the other six countries with significant publishing show a similar trend, although less pronounced than Brazil's (Figure 2). The first big increase occurs two years later, between 1999 and 2000 when SciELO began to promote itself more outside of its country of origin (Brazil). For example, there are 23 journals in 1996 from Mexico and 87 in 2000. The sharpest increase appears to be for the journals edited in Colombia when, the number of journals goes from 4 to 35 over the same time period.

The impact of Redalyc's appearance in 2005 is visible in all countries, especially Mexico and Colombia. The latter showed a sustained growth in the number of journals over the following decade, eventually surpassing Mexico, as the number of journals there slightly stagnated. Argentina, Venezuela, and Cuba follow a similar pattern with two sudden increases. However, the increase in subsequent years is moderate compared to others. More specifically, journals published in Cuba experienced an upward trend between 2005 and 2015 followed by a decline. The number of journals indexed in Venezuela decreased starting from 2009 and more rapidly from 2014. Finally, Chile's consistency sets it apart: 2005 did not imply a sudden change in the growth trend of journals.

Figure 2
Journals with at least one article published per year, by country of publication (excluding Brazil), 1995-2017

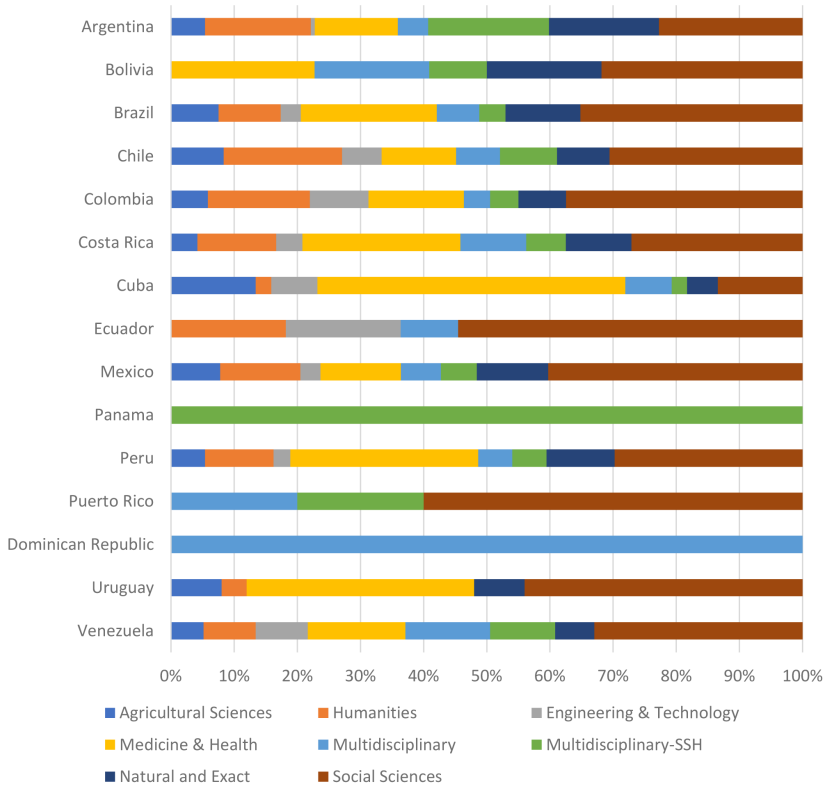


Source: The OLIVA database (June 2019). Note: journals from Bolivia, Costa Rica, Ecuador, Panama, Peru, Puerto Rico, the Dominica Republic, and Uruguay are included in "Others". The years 2018 and 2019 were excluded.

Let us now take a closer look at the disciplinary features of OLIVA database journals. Social Sciences journals represent 33% of the total, Medical and Health Sciences 19%, Humanities 12%, and Natural and Exact Sciences 10%. However, if we look at the disciplinary distribution by publishing country, we see that journals in the Social Sciences are predominant in

Uruguay, Ecuador, Puerto Rico, Brazil, Bolivia, Mexico, and Peru (see Figure 3). Medical and Health Sciences journals have an important share in Cuba, Uruguay, Bolivia, Costa Rica and Peru. Journals in the Natural and Exact Sciences are most present in Argentina, Bolivia and Brazil. Figure 3 shows the disciplinary areas of journals and makes it possible to compare their distribution in less visible countries, such as in Central America and the Caribbean, with the distribution in larger regional actors.

Figure 3
Journals by discipline and country of publication (n=1,720)



Source: The OLIVA database (June 2019).

The comparison by country is very interesting but it is also important to observe publication asymmetries within the countries. Scientific production is often centralized and concentrated in the main cities – particularly in the biggest institutions of higher education. For this reason, we are interested in surveying which institutions serve as the main journal publishers in Latin America and the Caribbean.

Among the 1,720 journals in the OLIVA database, we found 899 different publishers, which shows the diversity of the Latin American circuit. However, it is worth pointing out the institutions that publish a significant share of journals: 38% of the total are published by 53 mega universities, while the remaining 62% being published by 846 institutions. Table 2 shows the institutions that publish the largest number of journals.

Table 2
The 19 institutions that publish more than 10 journals

Publishing Institution	Number of journals published
Universidad Nacional Autónoma de México	61
Universidade de São Paulo, Brazil	47
Editorial de Ciencias Médicas, Cuba	38
Universidad de los Andes, Colombia	36
Universidad Nacional de Colombia	36
Universidad de Antioquia, Colombia	22
Pontificia Universidad Javeriana, Colombia	22
Universidad de Costa Rica	19
Universidad de Chile	16
Universidad de Buenos Aires, Argentina	16
Pontificia Universidad Católica de Chile	16
Universidade Estadual de Maringá, Brazil	**14**
Universidade Federal de Santa Catarina, Brazil	**13**
Universidad de Guadalajara, Mexico	12
Universidad Nacional de La Plata, Argentina	12
Universidade Federal de Rio de Janeiro, Brazil	11
Universidade Federal do Rio Grande do Sul, Brazil	11
Universidad del Valle, Colombia	10
Universidade Federal de Minas Gerais, Brazil	10
Total number of journals	**422**

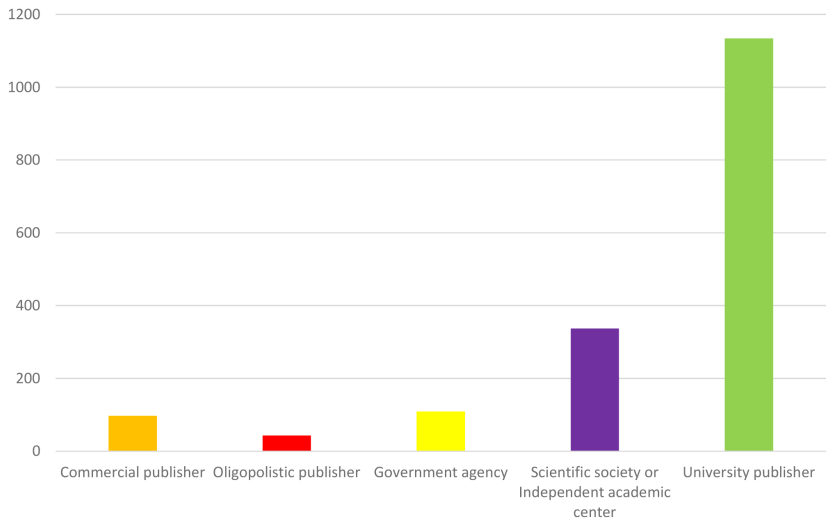
Source: The OLIVA database until June 2019.

As one may observe, the majority of journals featured in the survey are published by universities. The large public universities, such as the Universidad Nacional Autónoma de México, Universidade de São Paulo, Universidad de Costa Rica, Universidad de Chile and the Universidad de Buenos Aires bear considerable weight. Colombia is an exceptional case. Publishing in the country is concentrated in 5 institutions: Universidad de los Andes, Universidad Nacional de Colombia, Universidad de Antioquia, Pontificia Universidad Javeriana and Universidad del Valle. Another exceptional case is Cuba, where the Editorial de Ciencias Médicas (ECIMED) is the third largest institutional publisher of journals within the OLIVA database.[5]

Journal management and financing: The incidence of the Apc

Let us now look at the distribution of publishing institutions according by their editorial management and financing models. We found five types of publishers in the SciELO and Redalyc journals: 1) universities; 2) scientific societies, professional associations, independent research centers, national academies, and professional councils; 3) governmental agencies (ministry departments, museums, and publicly managed hospitals); 4) commercial publishers (here we gather small and medium size national and local specialized publishers), and finally; 5) big publishers. The final category refers to Reed-Elsevier, Wiley-Blackwell, Springer, and Taylor & Francis, considered to be oligopolistic commercial publishers as they concentrated between 50% and 70% of the world's scientific output until 2013 (Larivière Haustein, Mongeon, 2015).

Figure 4
Distribution of journals by type of publisher (n=1,720)



Source: The OLIVA database (June 2019).

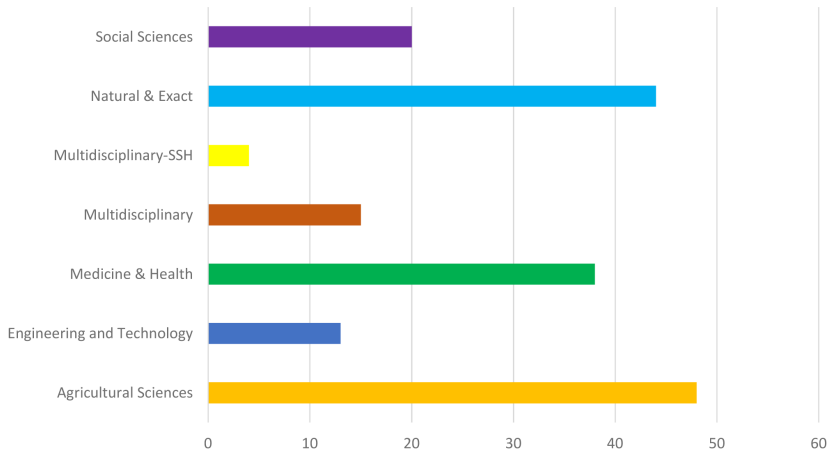
As Figure 4 shows, universities edit 66% of the journals, governmental agencies edit close to 7%, and scientific societies and independent research centers edit 20%. This shows a strong preference for academic and public management. We also identified and classified the commercial publishers' share: local or national commercial publishers (6%), and major commercial publishers (2.5%). Journals in medicine and health predominate within the local commercial publishers. Here is where ECIMED's share is noteworthy.

Major commercial publishers edit most of the journals they manage in Brazil (29).They also publish seven in Chile, three in Colombia, two in Mexico, and one in Argentina. Regarding discipline, this group (43) includes 20 journals that belong to Medical and Health Sciences, 14 to Exact and Natural Sciences, five to Engineering, one to Agricultural Sciences, and three to Social Sciences and Humanities. Forty-six percent of these journals are published entirely in English, and 88% of them belong to scientific societies that have outsourced their editorial management.

A key element in the current discussions of scholarly publishing refers to the change in the business model, from charging for subscriptions to charging fees to authors for publishing their papers (Article Processing Charges, APC). We gathered publication fee information from the websites of each of the journals. In most cases, journals did not state their respective publication fees. Upon further analysis, we verified that the absence of this information was because they were indeed open access diamond journals, i.e., they do not charge authors for publishing or readers for accessing their contents. In our universe of 1,720 journals, we observed that 10% of the journals charge APC's, most published in Brazil (60% of the total 182 journals with APC). This is different from what Appel and Albagli (2019) observed when analyzing a larger share of journals edited in Brazil that still showed a very low incidence of APC. In the OLIVA database, Brazil shows the biggest growth of this APC business model. To a lesser extent, in descending order, are the journals edited in Mexico, Chile, Argentina, Colombia, Peru, and Venezuela.

Figure 5 shows the journals with APC by discipline – 27% in Agricultural Sciences, 24% in Natural and Exact Sciences, and 21% in Medical and Health Sciences. This distribution shows the lower incidence of APC's among the Social Sciences that represent only 11% of the total (more than half of these belonging to Psychology). No journal with APC was detected in the Humanities, which is why this discipline does not appear in = Figure 5.

Figure 5
Journals with APC's by discipline (n:182)



Source: The OLIVA database (June 2019).

Among journals with APC's, scientific societies manage 38% of them, universities 37%, the major commercial publishers 13%, a governmental organization 7%, and small commercial publishers 5%. The fact that many journals managed by the academic community have sought to finance their operations through the adoption of APC's brings us to the question how sustainable the journals and their varying degrees of institutionalization are. Córdoba González (2021) observes that Latin American journals with APC's have largely low fees. This is also true for the journals within the OLIVA database, which suggests these are answers to the need to cover the costs of editorial processes and the management of the peer review process not funded by institutions. The journals indexed in SciELO and Redalyc have specific cataloging, evaluation and digitization procedures, which require material and human resources. Publishing institutions generally provide these, but in some cases, they turn to the APC to pay for the DOI, translations, or administrative or technical support. However, this cannot be equated to the business model of commercial journals. Furthermore, journals with high APC's, similar to the levels found in the mainstream industry, are uncommon in the region.

The sustainability of diamond access journals is a contested issue. The recently published OPERAS report highlights this model's importance, and it estimates the existence of 29,000 diamond access journals worldwide. It also points out their largely weak management as a generalized feature of these journals and suggests the adoption of a new business model (Bosman et al., 2021). The journals in SciELO and Redalyc have been described – the vast majority of which do not charge for reading or publishing – demonstrate just the opposite: that they *are not merely volunteer enterprises*. There is considerable consensus within Latin America about the need for national and institutional policies to sustain the management and regularity of the journals, but also a strong consensus that the solution does is not in the APC model (Babini and Debat, 2020). There is also a legitimate concern that the APC opens a floodgate to predatory journals and commercialization (Córdoba González, 2021).

Trends in collaboration and language of publication

Collaboration is an interesting and complex phenomenon in OLIVA journals. First, it should be noted that only 25.8% of the articles have a single author. This varies considerably among journals of the different

disciplinary areas (see Table 3). It represents 77.2% of the articles in the Humanities and 49.3% in Social Sciences. In the Humanities, only 19% of articles have two or three authors, and there are very few articles with more than three authors (3.8%). On the other hand, 40.8% of the articles in the Social Sciences have two or three authors, and 10% have four to eight authors. Table 3 shows that Multidisciplinary SSH journals have an intermediate percentage of authors compared to the two separate areas.

If the analysis is segmented by journals' country of publication, these trends within the social sciences undergo some variations. On average, journals edited in Brazil, Cuba and Puerto Rico have 38.6% of their articles written only one author, a rate significantly lower than that of the area as a whole. At the other end of the scale, Bolivia and Ecuador's Social Sciences journals are predominantly of single authorship (76.8% and 77.4%, respectively). The incidence of single authorship for journals edited in the remaining nine countries (Panama does not have any) ranges from 53.8% for Colombia to 62.4% for Mexico of articles with one author. Among the journals in the Humanities, differences are smaller. In Cuba, 64.2% of articles in Humanities journals have only one author. In Brazil, the percentage is at 70%, which is not far from Uruguay's 70.1% or Chile's 73.6% rates, for example. In Argentina, this rate rises above 80%, while in Mexico it reaches 87.6% and in Costa Rica 89.2%.

Table 3
Articles by journal discipline, by number of authors. (n=790,304)

Disciplinary Area	One author	2-3 authors	4-8 authors	More than 8 authors	Total
Agricultural Sciences	3.4%	27.6%	65.9%	3.1%	100%
Medical and Health Sciences	8.7%	30.7%	55.9%	4.7%	100%
Natural and Exact Sciences	11.5%	45.4%	40.7%	2.4%	100%
Engineering and Technology	8.0%	52.0%	39.2%	0.8%	100%
Multidisciplinary	17.7%	39.0%	40.2%	3.1%	100%
Social Sciences	49.3%	40.8%	9.7%	0.2%	100%
Humanities	77.2%	19.0%	3.7%	0.1%	100%
Multidisciplinary SSH	63.8%	29.7%	6.2%	0.2%	100%
All	25.8%	35.7%	36.2%	2.3%	100%

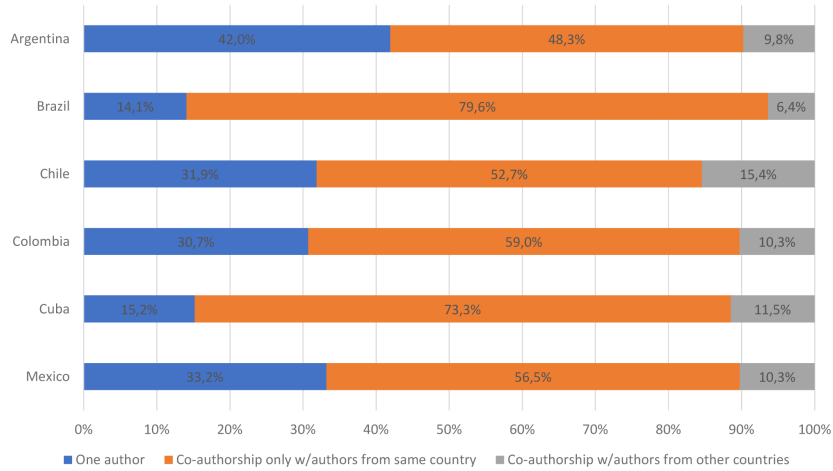
Source: The OLIVA database (June 2019).

Overall, the articles in the Agricultural Sciences, Medical and Health Sciences, and Multidisciplinary journals have mostly between 4 and 8 authors. Articles with two or three authors are located primarily in the Exact and Natural Sciences, and Engineering and Technology journals. In Multidisciplinary, they represent approximately 40% of articles.

Some countries stand out because their journals follow a different trend. In Brazil's Medical and Health Sciences journals, articles with only one author account for 5% of the articles, while those with four to eight authors represent 61.8% of the total. In all other areas, Brazil's journals show the same tendency of having articles with a greater number of authors. Another interesting case is Colombia, with a pronounced concentration of two or three authors per article (instead of 4 to 8 authors like the rest). In Agricultural Sciences, it totals 56.3% of articles, in Medical and Health Sciences 41.7%, in Natural and Exact Sciences 59.9%, and in Engineering and Technology 63.8%. All these data for Colombian journals show notably higher percentages than the aggregate of the rest of the countries in Table 3.

It is possible to observe the degree of international collaboration through co-authorship. Figure 6 selects six countries and shows the share of articles by a single author from the country in question, by several co-authors from that same country, and by co-authors from that same country with co-authors from other countries. It is interesting to compare the cases of Argentinian and Brazilian authors. A very high proportion – 42% – of articles authored by Argentinians are by a single person, a phenomenon related to the disciplinary dominance of SSH papers. The rest are split between those corresponding only to Argentine authors (48.3%) and those fruit of some form of international collaboration (9.8%). On the other hand, articles with only one Brazilian author are relatively few (14.1%). Here, collaboration among researchers affiliated to Brazil dominates (79.6%) and is distributed throughout the largest country in South America (see the next section). It is interesting to note that international collaboration is a very limited practice for Brazilian authors when they publish in Brazilian edited journals. Thus, the focus on national collaboration goes together with a lower propensity for single authorship, as already mentioned.

Figure 6
Articles with country affiliations in six selected countries, by type of authorship and type of collaboration. (n=492,308)



Source: The OLIVA database (June 2019). Note: only articles that mention the country of affiliation for all its authors are included.

It is interesting to compare these forms of collaboration OLIVA has been observing with other databases, where international collaboration between authors from these countries is much higher, and continues to grow. According to data from SCOPUS, 23.9% of articles authored by at least one Brazilian author in 2011 showed some form of international collaboration. This rate increased to 29.6% in 2015, and reached 36.1% in 2020. Argentina's rates these same years are 40.4%, 42.8% and 49.9%, respectively.[6] However, it is worth remembering that Brazil's share is significantly higher than that of the remaining countries in the region (according to SCOPUS, articles by Brazilian authors represent 48% of the Latin American documents produced between1996 and 2020).

It seems likely that the degree of international collaboration by Brazilian researchers in these mainstream journals is affected by the country's linguistic isolation within Latin America, its growing tendency to publish in English, and its internationalization policies that prioritize North America and Europe. Meanwhile, intra-regional collaboration is more common for Spanish-speaking Latin American authors. Based on the OLIVA corpus, one can say that Brazilian authors collaborate more with each other. Meanwhile, Argentines,

Mexicans, and Chileans work more regularly with colleagues from other countries in the region. However, it is important to mention that authors from Brazil represent an important share as co-authors in journals edited in other countries. For example, at least one author from Brazil features in 26.5% out of articles by authors from Argentina that count on international collaboration (n=3,906). In collaborative papers with authors from Chile (n=5,376), 19% of them feature co-authorship with Brazil. In Colombia's case (n=6,729), it is 16.6% of articles. Clearly, national collaboration for Brazilians seems to prevail when analyzing journals edited in Brazil, yet they participate significantly in the co-authorship of articles published in other countries. We will further explore the mapping of international collaboration in the OLIVA database in subsequent studies, especially to determine the balance between intra-regional and extra-regional collaboration, by discipline and by country. It will be worth exploring deeper whether authors from Brazil tend to collaborate with colleagues from outside the region more than with authors from Spanish-speaking countries, depending on the journal's recognition.

It is worth noting some nuances in the general trends with authors from Chile within the OLIVA database. The share of articles with either one single author or with international collaboration are more than double those of the Brazil's journals, with a lower incidence of collaboration within the country itself. Co-authorship with colleagues in other countries is the highest of all – a feature that, as explained below, may be indicative of a peculiar academic culture within the region. Among the remaining cases, Cuba shows to be similar to Brazil, with a greater degree of international collaboration. On the other hand, Colombia and Mexico are similar to Chile, but with a greater share of national collaboration.

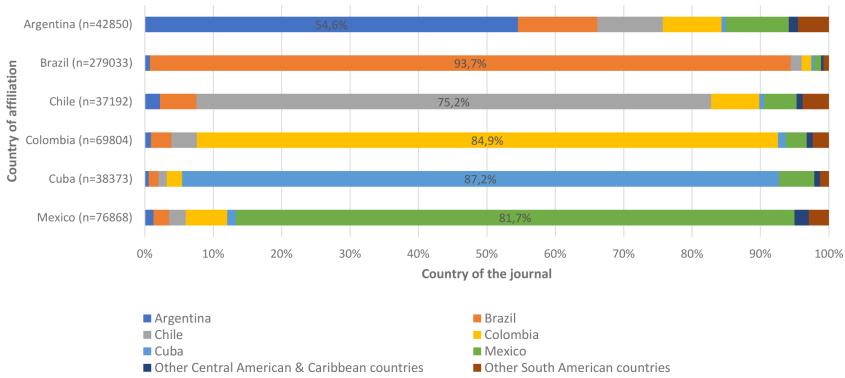
Along similar lines, by observing articles by authors of different countries and the journal's country, one can analyze trends analogous to those mentioned above (see Figure 7). The first thing that emerges from this analysis is that, for the countries in question, journals mainly publish articles by authors of their respective countries. This is especially true in the case of authors from Brazil: 93.7% of the articles they author are published in Brazilian journals. Chile comes second, yet it represents only 1.6% of total articles. In bibliometric studies based on the mainstream indexing sources, this characteristic is generally interpreted as academic inbreeding, thus frequently seen

in a negative light. In the following section, we propose an alternative approach, different from simply labelling national collaboration as academic inbreeding. This approach is based on a detailed study of the significant number of articles authored by Brazilians and published in Brazil.

Seventy-five percent of authors from Chile have published in national journals, which is certainly a high percentage but lower than others. Journals from Colombia, Brazil and Mexico are important regional recipients of Chilean authors' submissions. This rate may be linked to greater pressure from Chilean institutions' evaluation culture, which discourages publishing in nationally edited journals and provides direct incentives for international collaboration, as mentioned in the discussion of Figure 3. On the other hand, authors from Colombia, Cuba and Mexico tend to submit their work to nationally edited journals, although not to the same degree as seen in Brazil.

Argentina also has its peculiarities. The fact that it has the lowest share (54.6%) of articles published in national journals with at least one Argentinian author suggests a greater tendency to publish abroad than that observed in Chile. However, there are other explanations. There is evidence that journals from Argentina are underrepresented in the OLIVA database, due to the Latindex catalog's traditional selection of social science and humanities journals edited in Argentina. This selection is due to the nature of the existing evaluation culture at CONICET (the Argentine National Science Council) and national universities, where regulations reward publications indexed in Latindex at much as those in SciELO, Scopus and WoS, among other factors (Beigel, 2014). This is surely in addition to the lower technical complexity of indexing in the Latindex Catalog 2.0 compared to the SciELO and Redalyc requirements, and the weight of Latindex's national (Argentine) node. If we compare the Brazil, Argentina, Chile, Mexico and Colombia's journals in the Latindex Catalog 2.0 with the OLIVA database, we see that Chile has 144 journals in OLIVA and 167 in the Latindex Catalog 2.0. In contrast, there are 316 Argentine journals in the Latindex Catalog 2.0 and 167 in SciELO or Redalyc – almost twice as many. In contrast, Argentina's journals in the OLIVA database not belonging to the social sciences and humanities are mostly indexed in SciELO (43/69).

Figure 7
Articles by authors affiliated in the selected countries, by country of the journal



Source: The OLIVA database until June 2019. Note: The category "Other Central American and Caribbean countries" includes journals from Costa Rica, Panama, Puerto Rico and the Dominican Republic. "Other South American countries" includes journals from Bolivia, Ecuador, Peru, Uruguay and Venezuela. Each bar represents each country's total number of authors, while the colors in the bars represent the share published in journals of the corresponding countries.

Let us now move onto language trends. In contrast to the highly predominant role of English widely seen in mainstream databases, the OLIVA corpus is remarkably diverse in terms of language, driven by the infrastructure developed in the region with multilingual indexing and publishing protocols and infrastructures. Table 4 shows Spanish and Portuguese's central role, but also the significant presence of English that accounts for 23.9% of the articles. French has a very limited presence, as do other languages, some of which appear only once. Portuguese, logically, is found mainly in journals published in Brazil. It is interesting to note that 62.4% of the articles in Brazil's journals were published in Portuguese, while one-third (34.6%) was published in English. This value is only comparable to the percentage of English in articles in Puerto Rico (50.5%), Chile (23.7%) and Mexico's (17.1%) journals. Chile's rate of incidence might be another element pointing to an academic culture influenced more by *mainstream* standards. Mexico and Puerto Rico may be influenced by traditional academic ties.

Outside Brazil, Mexico, Puerto Rico and Chile's journals, an average of 5.4% of the articles are in English. On the other hand, Portuguese as a language of publication outside of Brazil's journals is very low: representing 5.4% of articles Costa Rica's journals, 3.8% of those journals edited

in Uruguay, and 2.5% in Ecuador. Conversely, in the Spanish-speaking area, Spanish is undoubtedly the predominant language. In journals from Bolivia, Cuba, Ecuador, Peru, Dominican Republic, Uruguay and Venezuela, Spanish accounts for more than 90% of articles.

Table 4
Articles by language (n=790,304)

Language	Articles [%]
Spanish	43.7%
Portuguese	32.09%
English	23.91%
French	0.2%
No data	0.12%
Total	100%

Source: The OLIVA database (June 2019).

For articles in journals edited in Brazil (n=396,293), Portuguese is most present in the Social Sciences and in Multidisciplinary-SS&H (85%). Articles in English are the majority in two disciplinary areas, namely Engineering and Technology (65.1%) and Natural and Exact Sciences (54.5%). Percentages are also high in Medical and Health Sciences (44.4%) and Multidisciplinary (42.4%). Spanish has a stronger presence than Portuguese in journals of Spanish-speaking countries. Consequently, 8.4% of the articles in the Multidisciplinary-SSH journals and 7.9% of those in the Humanities are in Spanish.

In journals from countries where Spanish is the official language, the share articles in English in Natural and Exact Sciences journals is 41.9% and for Engineering and Technology, 28.4%. Articles in Portuguese are very much a minority, and only in the Social Sciences does their rate reach 2.9%. Spanish, on the other hand, is dominant in the Medical and Health Sciences and in the three categories of Social Sciences and Humanities (more than 91%).

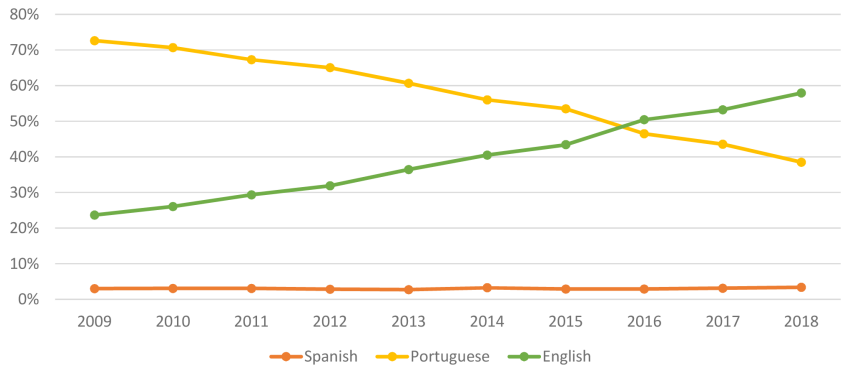
A greater number of authors tends to imply publications are in English. The average number of authors in English-language articles is 4.3. For Portuguese-language articles, the average is 3.3 and for Spanish, 2.7. This is partially due to the tendency of Brazilian authors and journals

to publish articles by two or more authors. The same decreasing order (English, Portuguese, and Spanish) in the average number of authors is present in all areas.

The authors' country affiliation is also affiliated to a publication being in English. The articles authored only by researchers from outside Latin America and the Caribbean (8.5% of 643,929 articles with country of affiliation known) represent 16.9% of the articles in English. Similarly, those in collaboration with at least one author from the region and one from outside the region account for 4.7% of the articles. In both, English is the main language of publication, both for articles with collaboration outside the region (49.9% of 30,307 articles) and for those without regional collaboration (52.4% of 54,700 articles).

It is also possible to analyze the evolution over time of language distribution among publications. In the last decade, English has increased its relative share at the expense of Portuguese, while Spanish has remained approximately constant (see Figure 8). A disaggregated examination of the two predominant languages – Portuguese and Spanish – reveals some trends that are worth highlighting. The share of articles in English among journals from Brazil rose from 24% in 2009 to 58% in 2018. The decrease in the relative weight of articles in Portuguese over this period was equivalent, while Spanish remained at a modest 3%. This trend of English is present in all disciplinary areas, although mild in the Social Sciences and Humanities, it is encouraged by the Agricultural Sciences, Medical and Health Sciences, Engineering and Technology, and Exact and Natural Sciences. The predominance of English was already noticeable in journals of the last disciplinary area at the beginning of the period analyzed.

Figure 8
Articles in journals from Brazil, by language and year of publication, 2009-2018 (N=790,304)



Source: The OLIVA database (June 2019). Note: the categories "other languages" and "no data" were excluded as they represent only 0.3% of articles.

In the other countries, although these ten years could also be characterized by a decrease in the relative weight of Spanish in favor of English, the latter never overtook the former. There is also an important disciplinary segmentation. In Agricultural Sciences, Engineering and Technologies, and Natural and Exact Sciences journals as a whole, English went from representing 29% of articles in 2009 to 42% in 2018. At the same time, Spanish dropped from 70% to 57% of articles. In other words, the official language of most Latin American countries remained the predominant language of articles.

On the other hand, the remaining disciplinary areas also show similar behavior. Overall, articles in Spanish in these journals dropped from 92.8% at the beginning of the selected period to 86.6% at the end of the period. In contrast, not only did the number of articles in English increase (5.6% to 10.2%) but also those in Portuguese (1.2% to 3%). This trend was slightly more present in Uruguay and Colombia.

National collaboration in Brazil

As previously mentioned, Brazil's weight is central to the analysis of the scientific journal space in Latin America. The country's journals account for 29.4% of the 1,720 journals in the OLIVA corpus. Analyzing these journals by discipline, most belong to the Medical and Health Sciences (34.3% of the total) and the Natural and Exact Sciences (33%). On the other hand, there are fewer among the Humanities (23.9%), Engineering and Technol-

ogy (19.3%) and Multidisciplinary-SSH (17.9%). However, Brazil's weight is best measured by analyzing the number of articles in its journals. Overall, 50.1% of OLIVA database articles have been published in Brazil's journals. By discipline, this percentage is even higher for journals in the Agricultural Sciences (65.8%), Multidisciplinary (58.8%), Medical and Health Sciences (54.5%), and Natural and Exact Sciences (53.9%).

This last section presents the analysis of a complementary data source to characterize national collaboration within Brazil in depth. It is the metadata of articles published in the journals of the SciELO Brazil collection over the 2016 to 2020 period, totaling 104,750 documents. This information was extracted from SciELO in May 2021. This share of the database has had author names and respective institutional affiliations standardized, which allows for a more precise analysis. This set of documents includes authors from 162 countries. In 74.5% of cases, there are only authors affiliated with Brazilian institutions (and 23.6% with only one author), while in 18.3% of cases there are only authors from other countries. In just 7.1% of the articles is there collaboration between authors inside and outside of Brazil. This paints a very similar picture to the one presented in Figure 6. Notably, in articles authored by Brazilians and foreigners, only 17% are co-authored with Latin American countries. This phenomenon is something that we have already commented on, and is largely due to linguistic aspects.

However, a heterogenous picture emerges when looking at the authors from Brazil in detail. There are articles by researchers from all 27 Brazilian states, but some concentrate a bigger percentage. In the set of documents with at least one author from Brazil with his/her geographic state declared (n=85.528), 85.7% belong to five states: São Paulo (33.4%), Minas Gerais (15.4%), Rio de Janeiro (13.7%), Rio Grande do Sul (13.1%) and Paraná (10.1%). In other words, states in the Southeast and South of Brazil.

Of this segment, 58,148 (63%) documents are co-authored. Among these, one can observe patterns of intra-national collaboration that moves the analysis beyond the dualism of academic inbreeding/cross-breeding. Indeed, 59% of the documents are by authors from the same state; however, 31.9% involve collaboration between authors from institutions in two different states. There is even 9.1% involving authors from three or more states (the maximum is 26 different states). This is an extremely complex and heterogeneous country, with a large number of university and scientific institutions, so that it is difficult to characterize a collaboration between two Brazilian authors (or publication in a journal of their

country) as academic inbreeding. It is possible to think along similar lines of collaboration between authors from different institutions within the same state.

In the absence of data to illustrate inter-institutional collaboration, the analysis of Brazil's interstate collaboration brings significant nuance to the interpretation of the forms of collaboration that appear in Figure 6. There we mention that 79.6% of the articles in the OLIVA database with authors from Brazil are exclusively by authors from Brazil, and that other countries show lower percentages of publication, such as Argentina (48.3%), Chile (52.7%) and Mexico (56.5%). The documents analyzed in this section (SciELO 2016-2020) show that, when authors from Brazil participate, there is collaboration in 41% of cases, either by authors from two states (31.9%) or from three or more states (9.1%). The remaining 59% (only intra-state collaboration) is similar to the rate of exclusively intra-national collaboration in the Spanish-speaking countries mentioned.

Indeed, collaboration is one of the fundamental factors in the remarkable growth of Brazilian authors' publications in recent decades. In this regard, Sidone, Haddad, & Mena-Chalco (2016) showed the central role of institutions in Southeastern states in intra- and inter-regional collaborations. Although these states continue to have this central role – particularly São Paulo – collaboration between other states in the same or different regions has increased in recent years. This phenomenon is particularly strong in the Northeast and South.

If one were to look at the scientific system of Brazil's magnitude at its highest level of aggregation, one would lose the detail needed to analyze the complexity of interactions inside its borders. Table 5 shows co-authorship by the number of different authors' states according to the disciplinary area of the journals (SciELO classification is used). Clearly, journals in disciplinary areas that tend to have a greater number of authors in their articles and documents exhibit greater interstate collaboration. Even the participation of authors from three or more states exceeds 10% in the Agricultural Sciences, Exact and Natural Sciences, Biological Sciences and Multidisciplinary journals. The degree of co-authorship in the three areas related to Social Sciences and Humanities is lower, but in no case does the number of documents with co-authorship between two or more Brazilian states fall below 29%. A significant portion of collaboration among authors is no longer between institutions, but between different states.

Table 5
Documents in SciELO-Brazil with two or more Brazilian authors, by interstate collaboration and journal discipline, 2016-2020 (N=58,148)

Disciplinary area	1 state	2 states	3 or more states	Total
Agricultural Sciences	48.3%	37.6%	14.1%	100%
Natural and Exact Sciences	50.2%	37.3%	12.5%	100%
Biological Sciences	51.6%	35.3%	13.1%	100%
Multidisciplinary	55.0%	33.3%	11.7%	100%
Engineering	62.0%	30.9%	7.1%	100%
Health Sciences	63.1%	28.4%	8.5%	100%
Humanities	64.8%	29.9%	5.3%	100%
Applied Social Sciences	66.2%	29.8%	4.0%	100%
Arts, Languages and Letters	68.9%	29.5%	1.6%	100%
All	59%	31.9%	9.1%	100%

Source: The SciELO Brazil database 2016-2020.

Nonetheless, there are evidently important asymmetries in the different variants of collaboration between authors from different states. If, for authors from a given state, one analyzes the other top five states with which the most papers are co-authored, São Paulo appears as a leading "partner" in 26 cases, Minas Gerais in 25, Rio de Janeiro in 21, Rio Grande do Sul in 14 and Paraná in 12. Pernambuco follows with seven, and Brasília with five. Seven states do not rank among the top five collaborative states for any other state.

Let us consider in detail the case of authors affiliated with institutions in the Northeast, for example. Table 6 shows which authors from other areas of the country have collaborated with each of the nine states of the region. First of all, there is a considerable tendency (just over 50%) for authors to collaborate with others from the same state. However, collaboration among the states of the region and the five most productive states (states with the most documents) as mentioned previously – São Paulo, Minas Gerais, Rio de Janeiro, Rio Grande do Sul and Paraná – is divided approximately equally. The total co-authorship with authors in the remaining 13 states ranges between 7.7% and 12.7%. In other words, collaboration "outside state borders" takes place with the same intensity

between states in the same region as it does between these same states and the five major centers of scientific production. Of course, the modest role of other states in intra-state collaboration adds to this trend.

Table 6
Documents in SciELO-Brazil with at least one author from the Northeast, by type of co-authorship

State	With co-authors from the same state.	With co-authors from other states in the Northeast.	With co-authors from the 5 most productive states.	With co-authors from other states in Brazil (13).	Total
Alagoas	42.9%	22.0%	22.6%	12.5%	100%
Bahía	47.3%	15.0%	27.4%	10.4%	100%
Ceará	48.4%	20.6%	21.6%	9.4%	100%
Maranhão	43.0%	16.4%	27.9%	12.7%	100%
Paraíba	45.2%	27.9%	19.2%	7.7%	100%
Pernambuco	46.9%	23.4%	20.9%	8.9%	100%
Piauí	43.4%	26.9%	20.3%	9.4%	100%
Rio Grande do Norte	46.7%	25.8%	19.2%	8.2%	100%
Sergipe	41.8%	23.0%	25.3%	9.9%	100%

Source: The SciELO-Brazil database, 2016-2020. Note: the "five most productive states" are those with the highest number of articles: São Paulo, Minas Gerais, Rio de Janeiro, Rio Grande do Sul and Paraná.

Another point to highlight is that the nine states analyzed exhibit similar percentages in each of the types of collaboration. The variations observed in the table are not significant, which may suggest that these are relatively stable features of the Brazilian scientific field. To support this conclusion, we look at a region that is very different from that of the Northeast. Table 7 shows the collaboration of the five most productive states. Its main feature is the degree of intrastate collaboration, yet higher than the Northeast – since the rate exceeds 50% in all cases. For each state in the table, the other four states represent approximately the same share of collaboration (fourth column as all the remaining states in Brazil). Clearly, the concentration in collaboration within these five states' institutions stands out. However, it is also clear that even though smaller, their collaboration with states less prominent in the Brazilian scientific sphere is not totally marginal. Finally, it is worth noting the

remarkable similarity between the profiles of the five states in Table 7, which might be suggesting – as aforementioned – consistent trends in this field.

Table 7
Documents in SciELO-Brazil with at least one author from one of the five states with the most documents produced, according to type of co-authorship

State	With co-authors from the same state.	With co-authors from the 5 most productive states. (other).	With co-authors from the rest of Brazil (22)	Total
Minas Gerais	52.9%	24.4%	22.7%	100%
Paraná	53.2%	26.6%	20.2%	100%
Rio de Janeiro	54.4%	23.7%	21.9%	100%
Rio Grande do Sul	58.4%	20.5%	21.1%	100%
São Paulo	57.9%	20.1%	22.0%	100%

Source: The SciELO-Brazil database between 2016 and 2020.

This analysis of collaboration within Brazil in recent years has highlighted the fact that low international collaboration in co-authored publications of a country does not imply institutional inbreeding or parochialism. Collaboration between authors from different states is a feature of the Brazilian scientific output space as analyzed through the SciELO documents of the the period between 2016 and 2020. This is a question of a practice of collaboration between teams located in clearly different spaces, with a diverse academic history and processes of institutionalization that affect particular disciplinary areas. This is because, as one may expect, interstate collaboration is less prevalent among disciplinary areas that tend to have fewer authors in their articles.

However, the concentration of authors responsible for these papers in Southeastern and Southern institutions does not dismiss the significant collaboration with colleagues from the rest of the country's extensive geography. States such as São Paulo, Minas Gerais or Rio de Janeiro can hardly be considered as homogeneous within one another. They include diverse institutions – federal, state, private – that presumably also collaborate on publications. This makes the idea of academic inbreeding in terms of authors' intra-state affiliations more nuanced given that collaboration

between authors from Unicamp (Universidade Estadual de Campinas), for example with those of USP (Universidade de São Paulo) – both in the same state – can hardly be classified as provincial interaction.

Conclusions

Many experts warned in the 1980s and 1990s that Latin American scientific output was not visible to the world as it was rarely cited in "international" databases, which is why Gibbs (1995) spoke of a "lost science of the third world." In fact, in those decades the region strengthened an infrastructure for scholarly communication, which sought to raise the value of the region's research output, with its unique disciplines, research agendas, and language of publication. This encouraged policies and the regional management of the circulation of Latin American science, giving rise to a strong and consolidated structure of journals, indexing bases, institutional repositories, networks of regional institutes, and large national universities that became regular publishers of academic journals. This arrangement grew fragmented as, despite strenuousinstitutional efforts, interoperable systems have not yet prevailed to allow for Latin American scientific output to be known, visible and valued worldwide. With this objective in mind, the OLIVA project was developed at the Universidad Nacional de Cuyo (Mendoza, Argentina), to enhance the value of the output assessed and indexed in articles to grow its presence and impact in academic evaluation systems.

The corpus of documents analyzed in this work is the result of a database that combines information from two of the main indexers of Latin American and Caribbean scientific journals – SciELO and Redalyc. Thanks to the collaborative work of both institutions and CLACSO, we can appreciate the decades of effort made by these open access journal platforms and the extent to which there is high quality Latin American output whose value is still underrepresented in current academic evaluation models. Perhaps one of the main findings emerging from the analysis is the longevity and diversity of journals, some of which were created in the late 19th and early 20th centuries.[7] Diverse disciplinary areas are also represented, as well as forms of collaboration, languages of publication, types of authorship, authors' geographic state of affiliation, and publishing institutions over a period of more than a century of Latin American and Caribbean scientific production.

The analysis of collaboration structures in a corpus such as OLIVA delved into the complexities of national collaboration to distinguish between co-authorships belonging to the same institution and collaborations between different universities and especially those from different states, thereby examining the geospatial dynamics of national scientific output. Brazil's analysis, which represents one-third of the total number of journals as well as half of the authors and articles in the OLIVA database, showed that there is strong collaboration between researchers from different states throughout the country. We also characterized the states that play an important role in the corpus on their own, and yet sseem to collaborate significantly with colleagues from states that are less significant within the Brazilian scientific system.

An examination of the publishing institutions of OLIVA corpus journals shows structural differences with mainstream commercial journals, such as in the degrees of autonomy and attachment to the academic community. Neubert and Rodrigues (2021) analyzed Latin American output in the Web of Science and observed that these articles were published in 11,965 scientific journals, with commercial publishers accounting for 56.48%, universities for 18.30%, and associations for 19.25%. On the other hand, SciELO and Redalyc journals, are extremely academic and published almost entirely by universities, scientific societies and government agencies, offering a platform of editorial services anchored in the public domain. The region's scholarly community has historically been in charge of creating journals, and they continue to support them materially even though evaluation systems encourage their academics to publish in mainstream journals.

Another intrinsic feature of scientific publication in Latin America is the widespread adoption of Open Access policies, nurtured by the decision of research funding agencies, universities, scientific societies, and publishers to provide non-profit infrastructures for visibility in the global flow of scientific information. It is worth noting that the entire OLIVA corpus is available in open access and full text on the SciELO Network and Redalyc platforms, as well as on each journal's website, and in thematic and institutional repositories. In addition to the output analyzed in these 1,720 journals, there are thousands of other journals in the region that show evidence of long-term projects that have been key to sustaining open access in the region.

The study of the output published in SciELO and Redalyc has allowed us to discuss some assumptions that permeate the recognition of scientific knowledge. In other words, prioritizing so-called "mainstream" journals while discrediting the rest. OLIVA reveals two dimensions of this scientific communication space of growing editorial quality. First, its diversity, and thematic and linguistic breadth. Second, the international and inter-institutional collaboration at the national level, which encourages a perspective beyond the traditional duality that opposed the "mainstream or global circuit" – led by the most prestigious journals – and the "regional or local circuits," with limited circulation. This duality, promoted in a period of commercialization of journals and reinforced by evaluation systems through journals' Impact Factor, has undervalued non-commercial open access publications such as those we are concerned with in this study. This perspective also fails to appreciate the forms of intra-national collaboration, whose complexity we seek to show in this paper in Brazil's case, and which are particularly relevant for research addressing local issues that are enhanced in this type of collaboration. We also point out the linguistic particularities of Brazil, which largely explain trends in collaboration when nationally edited journals are up against mainstream journals.

The forms of collaboration shown by the journals indexed in SciELO and Redalyc and their academic and university anchoring show that this system promotes broad dissemination and multi-scale research programs. It can very effectively address the current needs of scientific communication in times of open science. This is not only based on Latin American professional scientific publishing, indexing and open access, but also on its experience in producing interoperable regional indicators as seen in the joint publication developed with the main indexing systems in the region, the PKP and CLACSO (Alperin, Babini and Fischmann, 2014). More efforts to improve the visibility and comparability of Latin American output will enable the region to counteract a major trend in two hitherto separate but increasingly intrinsically linked forces: commercial threats to the open science movement and the urgent need to reform traditional forms of research evaluation. It is our hope that the OLIVA project will stimulate more collaborative spaces to promote a socially significant and participatory science.

(Received on July 15, 2021)
(Resubmitted on January 27, 2022)
(Approved for publication on April 7, 2022)

Notes

1. https://pkp.sfu.ca/ojs/ojs-usage/ojs-stats/
2. Latindex Catalog 2.0. is a comprehensive list of scholarly journals published in the Ibero-American region which are individually assessed to guarantee a set of quality criteria indicators. The Latindex portal shows information of the journals but not at the level of document.
3. See https://cecic.fcp.uncuyo.edu.ar/oliva/
4. SciELO also includes journals from Spain, Portugal, and South Africa while Redalyc includes journals from Poland, Germany, Spain, Portugal, India, and Angola. In the meantime, OLIVA is building an additional corpus with Spain and Portugal.
5. ECIMED was founded in Cuba in 1998 as the publishing house of the Centro Nacional de Ciencias Médicas that by then already edited 15 journals. The Center was part of the development of the medical sciences the Cuban government promoted. ECIMED also publishes books and its official website functions as an open access portal for the 38 journals available in full-text.
6. Values obtained by searching in SCImago on July 1, 2021 (https://www.scimagojr.com/countrysearch.php?country=BR).
7. Memórias, created in 1909, is the journal with the oldest articles included in OLIVA. Meanwhile, there are other journals indexed in SciELO that were created in the 19th century, such as: *Gaceta médica de Méxic*o (1864), *Gaceta Médica de Caracas* (1893), *Revista chilena de historia natural* (1897), *Revista de Ciências Agrárias* (1903), *Boletín de la Sociedad Geológica Mexicana* (1904), *Ingeniería, investigación y tecnología* (1908), *Memórias do Instituto Oswaldo Cruz* (1909), *Revista industrial y agrícola de Tucumán* (1910).

References

ALPERIN, J. P., Babini, D.; Fischman, G. (Eds.) 2014. Indicadores de acceso abierto y comunicaciones académicas en América Latina. Buenos Aires: CLACSO. Retrieved from http://biblioteca.clacso.edu.ar/clacso/se/20141217052547/Indicadores_de_acceso_abierto.pdf

APPEL, A. L.; Albagli, S. (2019). The adoption of Article Processing Charges as a business model by Brazilian Open Access journals. *Transinformação*, 31, e180045. https://doi.org/10.1590/2318--0889201931e180045

BASSON, I., Simard, M.A., Aubierge Ouangré, z., Sugimoto, C.R.., Lariviére, V. 2021. Data sources and their effects on the measurement of open access. Comparing dimensions with the web of science. In, Proceedings of the 18th International Conference on Scientometrics and Informetrics. pp. 93-98, on line.

BEIGEL, Fernanda. 2019. "Indicadores de circulación: una perspectiva multi-escalar para medir la producción científico-tecnológica latinoamericana". *Ciencia, Tecnología y Política* 2(3):028. doi: 10.24215/26183188e028.

BEIGEL, F. 2014 "Publishing from the Periphery: Structural heterogeneity and segmented circuits. The evaluation of scientific publications for tenure in Argentina's CONICET". *Current Sociology*, v. 62 n. 5, pp. 743–765.

Bosman, Jeroen, Jan Erik Frantsvåg, Bianca Kramer, Pierre-Carl Langlais, and Vanessa Proudman. 2021. *OA Diamond Journals Study. Part 1: Findings*. Zenodo.

Chittó Stumpf, Ida Regina, Samile Andrea de Souza Vanz, Ana Maria Mielniczuk de Moura, and Sônia Elisa Caregnato. 2017. "Scientific Output Indicators and Collaboration in Southern Brazil". *Revista Interamericana de Bibliotecología* 40(1):45–57. doi: 10.17533/udea.rib.v40n1a05.

Córdoba González, S. (2021). Cobrar por publicar en revistas académicas, una amenza al ecosistema latinoamericano no comercial, en Becerril-García, A. y Córdoba González, S. Eds. Conocimiento abierto en América Latina: trayectoria y desafíos. Ciudad Autónoma de Buenos Aires: CLACSO, pp.175-202.

Gibbs, Wayt (1995). Lost Science in the Third World. Scientific American, v. 273, n. 3, pp. 92-99.

DOI: 10.1038/scientificamerican0895-92

Gingras, Yves. 2016. *Bibliometrics and Research Evaluation: Uses and Abuses*. Cambridge, MA, USA: MIT Press.

Guédon, JC (2011) El acceso abierto y la división entre ciencia "principal" y "periférica". Crítica y Emancipación, v. III, n. 6, pp. 135-180.

Larivière V, Haustein S, Mongeon P (2015) The Oligopoly of Academic Publishers in the Digital Era. PLoS ONE 10(6): e0127502. https://doi.org/10.1371/journal.pone.0127502

Martinovich, V. 2020 "citation indicators and scientific relevance: Genealogy of a representation". *DADOS*, Rio de Janeiro, v. 63(2): e20190094.

Miguel, S. 2011 Revistas y producción científica de América Latina y el Caribe: su visibilidad en SciELO, RedALyC y SCOPUS. Rev. Interam. Bibliot [online] v. 34, n. 2, pp. 187-198.

Neubert, Patricia da Silva, and Rosângela Schwarz Rodrigues. 2021. "Oligopólios e publicação científica: a busca por impacto na América Latina". *Transinformação* 33: e200069. doi: 10.1590/2318-0889202133e200069.

Packer, Abel Laerte. 2005. "A construção coletiva da Biblioteca Virtual em Saúde". *Interface - Comunicação, Saúde, Educação* 9(17):249–72. doi: 10.1590/S1414-32832005000200004.

Packer, Abel Laerte. 2020. "The Pasts, Presents, and Futures of SciELO". en *Reassembling Scholarly Communications: Histories, Infrastructures, and Global Politics of Open Access*, editado por M. P. Eve y J. Gray. The MIT Press.

Packer, Abel Laerte, and Rogerio Meneghini. 2007. "Learning to Communicate Science in Developing Countries". *Interciencia* 32(9):643–47.

Ràfols, Ismael. 2019. "S&T Indicators in the Wild: Contextualization and Participation for Responsible Metrics". *Research Evaluation* 28(1):7–22. doi: 10.1093/reseval/rvy030.

Rodríguez García, Ariel, ed. 2020. *La revolución de los datos bibliográficos, científicos y culturales*. México: UNAM, Instituto de Investigaciones Bibliotecológicas y de la Información.

Salatino, Maximiliano, and Banzato, Guillermo. 2020. "Confines históricos del acceso abierto latinoamericano". in *Conocimiento Abierto en América Latina: trayectoria y desafíos*, edited by A. Becerril García y S. Córdoba. In print.

Sidone, Otávio José Guerci, Eduardo Amaral Haddad, and Jesús Pascual Mena-Chalco. 2016. "A ciência nas regiões brasileiras: evolução da produção e das redes de colaboração científica". *Transinformação* 28(1):15–32. doi: 10.1590/2318-08892016002800002.

Vessuri, Hebe, Jean-Claude Guédon, and Ana María Cetto. 2014. "Excellence or Quality? Impact of the Current Competition Regime on Science and Scientific Publishing in Latin America and Its Implications for Development". *Current Sociology* 62(5):647–65. doi: 10.1177/0011392113512839.

Vessuri, Hebe M. C. 1994. "La ciencia académica en América Latina en el siglo XX". *Redes : revista de estudios sociales de la ciencia* 1(2), 41-76.

Vuotto, Andrés, Victoria Di Césare, and Natalia Pallotta. 2020. "Fortalezas y debilidades de las principales bases de datos de información científica desde una perspectiva bibliométrica". *Palabra Clave (La Plata)* 10(1):e101. doi: 10.24215/18539912e101.