Fernanda Beigel, Abel L. Packer, Osvaldo Gallardo, Maximiliano Salatino

https://doi.org/10.1590/SciELOPreprints.4637

Fernanda Beigel
PhD in Political and Social Sciences. Principal Researcher at CONICET. 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
https://orcid.org/0000-0002-7996-9660

Abel L. Packer
Master of Library Science, Syracuse University. Project Coordinator for the Fundação de Apoio à Universidade Federal de São Paulo, Director of the SciELO / FAPESP (Scientific Electronic Library Online) Program.
Email: abel.packer@scielo.org
https://orcid.org/0000-0001-9610-5728

Osvaldo Gallardo
PhD in Latin American Social Studies, Universidad Nacional de Córdoba. Postdoctoral Fellow at the Agencia Nacional de Promoción Científica y Tecnológica. 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
https://orcid.org/0000-0003-0662-2196

Maximiliano Salatino
Email: maxisalatino@gmail.com
https://orcid.org/0000-0001-5573-9599
ABSTRACT
This article presents the results of the Latin American Observatory of Evaluation Indicators (OLIVA, its Spanish acronym) which works to raise the visibility of the scientific production indexed in Latin America and the Caribbean, and to enhance its value for research assessment. This study deals with the production published in open access by journals indexed in SciELO and Redalyc, based on a deduplicated database of a total of 908,982 documents and 2,591,704 authors. It highlights the magnitude of this production and describes its disciplinary diversity, as well as trends in national, regional, and international research collaboration. It also examines the publishers of the 1,720 journals that make up the database of journals from 15 countries and highlights the predominance of universities and public institutions in this regional circuit. The journals that operate with the APC model are analyzed, confirming a clearly lower influence of this model compared to what is observed in other continents, although Brazil is identified as the country with the greatest incidence of this practice in the region. Finally, collaboration between states in Brazil, analyzed with SciELO data, proves to be highly significant, strongly challenging the traditional interpretation of co-authorship among researchers from the same country as academic inbreeding. The study concludes that these journals show multi-scalar circulation, a linguistic diversity, and a disciplinary breadth that can very effectively serve the current needs of scholarly communication in times of open science.

Keywords: Scholarly communication; SciELO; Redalyc; National collaboration; International collaboration.


RESUMO
Este artigo apresenta os resultados do Observatório Latinoamericano de Indicadores de Avaliação (OLIVA) que visa contribuir para tornar mais visível a produção científica indexada na América Latina e Caribe, a fim de promover a sua valorização nos sistemas de avaliação. Este trabalho aborda a produção científica em acesso aberto das revistas indexadas no SciELO e Redalyc, com base na consolidação de uma base de dados, sem sobreposição, com o número total de 908,982 documentos e 2.591.704 autores e autoras. Também analisa as instituições editoras das 1.720 revistas de 15 países que compõem a base de dados, destacando a predominância das universidades e instituições públicas no circuito regional. O estudo destaca a magnitude desta produção, a sua diversidade disciplinar e as tendências da colaboração nacional e internacional. As revistas que operam sob o modelo APC são delineadas, o que confirma uma incidência muito inferior à de outros continentes, embora o Brasil seja apresentado como o país com o maior peso desta prática. Finalmente, a colaboração entre estados do Brasil, analisada com dados da SciELO, se revela altamente significativa, confrontando a interpretação tradicional de coautoria entre pesquisadores do mesmo país em termos de endogamia. O estudo conclui que estas revistas mostram formas de circulação multiescalar, diversidade linguística e amplitude disciplinar que podem contribuir de forma muito eficaz para as necessidades atuais de comunicação científica em tempos de ciência aberta.

Palavras-chave: publicação científica; SciELO; Redalyc; colaboração nacional; colaboração internacional.

RESUMEN
El artículo presenta los resultados del Observatorio Latinoamericano de Indicadores de Evaluación (OLIVA) que se propone contribuir a visibilizar la producción científica indexada en América Latina y el Caribe, para promover su valorización en los sistemas de evaluación. En este trabajo se aborda la producción publicada en acceso abierto por revistas indexadas en SciELO y Redalyc, a partir de la consolidación de una base de datos, sin solapamientos, con un total de 908.982 documentos y 2.591.704 autores y autoras. El estudio destaca la magnitud de esta producción y describe su diversidad disciplinar, así como las tendencias en la colaboración nacional, regional e internacional. Se analizan también las instituciones editoras de las 1.720 revistas que componen la base de datos, pertenecientes a 15 países, destacando el predominio de las universidades y las instituciones públicas en este circuito regional de publicación. Se detallan las revistas que funcionan con el modelo de APC, lo que confirma una incidencia francamente menor a la que se registra en otros continentes, aunque Brasil se presenta como el país con mayor peso de esta práctica. Finalmente, la colaboración interna entre estados de Brasil, analizada con datos de SciELO, se revela altamente significativa confrontando la lectura en términos de endogamia que tradicionalmente se ha asignado a la co-autoría entre investigadores/as de un mismo país. El estudio concluye planteadando que en estas revistas se observan formas de circulación multi-escalar, una diversidad lingüística y una amplitud disciplinar que pueden contribuir de un modo muy eficaz a las necesidades actuales de la comunicación científica en tiempos de ciencia abierta.

Palabras clave: publicación científica; SciELO; Redalyc; colaboración nacional; colaboración internacional

Conflict of Interest
The authors declare that there is no conflict of interest.

Funding
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).

CRediT -Author's contribution Statement (extracted from DADOS – Scholar One)

1. Beigel, Fernanda: Conceptualization (Equal); Data curation (Equal); Funding acquisition (Lead); Investigation (Equal); Methodology (Equal); Project administration (Lead); Resources (Equal); Validation (Equal); Visualization (Equal); Writing –original draft (Equal); Writing – review & editing (Equal).
2. Packer, Abel: Conceptualization (Equal); Data curation (Equal); Validation (Equal); Visualization (Equal); Writing –original draft (Equal); Writing –review & editing (Equal).
3. Gallardo, Osvaldo: Conceptualization (Equal); Data curation (Equal); Investigation (Equal); Methodology (Equal); Validation (Equal); Visualization (Equal); Writing –original draft (Equal); Writing –review & editing (Equal).
4. Salatino, Maximiliano: Conceptualization (Equal); Investigation (Equal); Methodology (Equal); Validation (Equal); Visualization (Equal); Writing –original draft (Equal); Writing –review & editing (Equal).

DADOS, Rio de Janeiro, e20210134 (Versão Preprint)
1. Introduction

The Latin American academic circuit is a constellation of national academic communities and regional networks that was formed in the mid-20th century with the support of inter-governmental organizations, cooperation agencies and foundations that promoted the management of scientific information as a key element for development. This led to the emergence of regional institutions that collaborated closely in the cataloging and construction of bibliographic indexes to foster the dissemination of locally produced scientific knowledge (Rodriguez Garcia Coord. 2020). By 1967, a significant step forward was made by the building of information centers such as BIREME and research networks such as CLACSO. They soon became major regional repositories for the dissemination of Latin American scientific production (Packer, 2005; Vessuri, 1994). Efforts to professionalize scientific publishing were driven by the creation of indexing systems led by regional centers affiliated to 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 afterwards Latindex and BIBLAT). All these systems were created and operated by universities, national and inter-governmental organizations to improve the bibliographic exchange of the region’s scientific and technical production. 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. In addition, and in conjunction with Latindex, these services created a journal quality certification system that focuses on peer review and best practices in editing. This 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 the development of open science (Vessuri, Cetto and Guédon, 2014; Beigel 2019; Packer, 2020; Banzato and 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 for reading, and points to the role of Latin America in the production of 25% of the total of these publications worldwide. Indeed, these open access journals are mostly published by universities and managed by full-time professors, assisted by technical support teams centralized in the libraries. An important aspect in the consolidation of university publishing in this region is the massive adoption of the open source Open Journal System (OJS-PKP) for the management of the editing, peer review and publication of the journals. In addition to university journals, there are also those of the scientific societies and professional associations which are often operated with the support of university infrastructure and personnel for digitization, creating permalinks, acquiring DOI’s, and doing the XML markup of the texts. Although this support needs to be strengthened, this core institutional and governmental support explains the existence and growth of Latin American and Caribbean journals.

---

1 We are deeply grateful for the relevant comments received by the two peer reviewers of this article, and for the thoughtful review by Dominique Babini.
2 https://pkp.sfu.ca/ojs/ojs-usage/ojs-stats/
The vibrancy of this regional circuit is now more visible globally 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 has a greater coverage of Latin American and Caribbean production. However, there is still no significant impact of these journals on global S&T reports or on evaluations of scientific careers, projects and institutions in which the use of the Impact Factor and other indicators from traditional databases such as the Web of Science and Scopus continue to prevail. This prejudices the value of the journals indexed in the region because it drives many research groups to preferentially choose journals managed by the dominant players in the publishing industry. This choice has been historically shaped in the academic community by the deeply rooted belief of the relevance/transparency of these indicators (Martinovich, 2020). It also highlights the gap between the ability of the region to, on the one hand, produce and publish quality science and, on the other, to gain international legitimacy for its journals (Packer and Meneghini, 2007). Several studies have already pointed out the limitations of these traditional indicators which were favored by several structural conditions and generated a series of abuses and distortions that affected the creativity and even the researchers’ concern for the social relevance of science (Guédon, 2011; Gingras, 2016; Ráfols, 2019). Thus, policies of internationalization that are justified by the need to advance in the rankings result in many institutions having their autonomy restricted and a weakening of their own capacities to link with the research environment.

The development of a global consensus on the need for open access to scientific literature has brought increasing attention to the region’s journals and indexing systems. Basson et. al (2021) analyze the proportion of open access articles available in Dimensions, which by 2021 had 46.6% of its documents in open access. They note that this percentage increases to 63.5% for documents published by Latin American and Caribbean journals. There are few studies that provide information, at the document level, on the breadth and diversity of the production published in the region, apart from the aforementioned mainstream databases. An important precedent is the study by Miguel (2011) which compared the coverage of journals from Latin America and the Caribbean included in SciELO, Redalyc and SCOPUS for the period 2005-2009. The estimated volume of scientific production using these three sources revealed differences with respect to the coverage registered in the Latindex catalog3. In relation to disciplinary coverage, SCOPUS and SciELO were more thematically balanced, while Redalyc showed a greater inclination towards the Social Sciences and Humanities.

Vuotto, Di Césare and Pallotta (2020) analyzed 17 bibliographic databases and verified that currently there are still few of them that openly offer their elements in a common format and structure to enable the analysis of production at the document level. For this reason, to date, the available empirical studies have only been able to analyze the journal collections separately as offered by each of the indexing databases. Scopus has 890 journals that are published in the region, and Web of Science has 223. As of June 2021, SciELO had 1,358 active, indexed journals which are also replicated in the SciELO Citation Index in the Web of Science platform, allowing for the evaluation of the performance of the articles in a global context of citations. Redalyc, for its part, has its own information system and indicators for its collection of 1,415 journals. These two collections have many

---

3 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.
journals in common, as we will see shortly. Then there is also Biblat which has its own cataloging system with a large number of complete records at the document level. However, at present it is not possible to work with combined data from these three indexing databases in the region because they are not interoperable. To these sources LA Referencia should be added. It is a very relevant database which harvests the production of 790 institutions from 12 Ibero-American countries. It currently has more than three million documents. However, it does not have advanced searching by journal, and there are overlaps of data with the former indexing services given that it originates from the repositories of the individual institutions.

Furthermore, there are thousands of journals that do not have metadata at the document level. This affects the possibility of analyzing the production published in Latindex journals which continues to be, as Miguel already observed in 2011, the cataloging database with the greatest coverage. In addition, there is no standardization of institutional affiliations, nor of 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 different databases. All these further limit studies at the document level. Consequently, this large corpus of documents published in the region is observed from fragmented approaches.

It is in this context that the OLIVA project (Latin American Observatory of eVAluation Indicators) emerged as an initiative of the Centro de Estudios de la Circulación del Conocimiento (CECIC-Facultad de Ciencias Políticas y Sociales), which was approved and funded in early 2019 by the Universidad Nacional de Cuyo. The main purpose of this project is to give visibility to the richness of this scientific production published in the region, and to develop indicators to recognize the value of these journals for the assessment of 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. The results of this first phase of the project are presented in this paper. The second phase foresees the incorporation of documents from journals indexed in Latindex Catalog 2.0 and in Biblat. A pilot is already underway with these institutions.

A database of 1,720 scientific journals published in Latin America and the Caribbean and indexed in SciELO and/or Redalyc was built within the framework of OLIVA. The study was organized on two levels - one at the journal level, and the other 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 on the description of the journals in the database constructed at this stage (SciELO and Redalyc). We then pass on to characterizing the publishing institutions of the journals and highlight their predominantly public nature and self-management by the academic community. An empirical classification of the journals’ disciplines was carried out based on the information found on their respective web sites. This was done because of the existing limitations of classifying the documents based on the disciplinary cataloging offered by the indexing databases. It revealed that the production has broad coverage, with half of the indexed documents being from the Social Sciences and Humanities, and the other half from Engineering, Natural Sciences,

---

4See https://ce Tic.fcp.uncuyo.edu.ar/oliva/
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 that publish in English and a multilingualism that is emerging in journals that are beginning to publish in several languages.

Finally, we focus on the case of Brazil. Given the significant weight of this country with 50% of the total number of articles published in the database and having a high tendency towards co-authorship among Brazilian 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 towards scientific collaboration between authors affiliated with the same country but belonging to universities from different states. A picture thus emerges that calls into question the preconception that collaboration between authors of the same country signifies academic inbreeding and reveals the complexity of this national academic space and its current dynamics.

A note about the sources and methods

The OLIVA project database is made up of the collected data of the journals indexed by SciELO and Redalyc, and revised data from the published documents. Each of these indexing systems collaborated by providing its historical database updated to June 2019. The data was subjected to data cleaning and deduplication of the journals in common. 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 as this is outside the scope of the project.

The construction of this integrated database faced significant limitations. The principal one was, as has already been pointed out, the lack of a common identification system amongst the 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 diverse portions of the indexed published documents in the journals pertaining to both indexing services.

In view of all this, we opted for a process of revision of the journals through a manual and meticulous procedure. First, the journals only in SciELO or only in Redalyc were identified, and the metadata of these documents were aggregated including articles, reviews, and editorials - both from active journals at the time of capture as well as documents indexed at an earlier time in between the gaps. In coverage. As for the journals common to both indexing platforms, the same journal could have different gaps or periods of coverage in each of the platforms. Thus, the consistency and completeness of the data were compared and, for each case, it was decided to keep those that had the widest temporal coverage and the largest number of documents. It should be clarified that in this phase of the project, we worked only with journals published in Latin America and the Caribbean.

Regarding the disciplinary classification of journals, we faced the challenge of the different classifications used by the indexing databases, even for the grouping into scientific areas. Thus, the same journal could be classified in different disciplines and, as a result, in different areas. For the OLIVA database, we used a classification based on the OECD and Frascati Manual criteria. The assignment of one or more disciplines to each journal was made through an empirical survey, based on the description that each journal provides in the "Scope" or "About" section of its official website. Then the journals were 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; 7) Multidisciplinary; and 8) Multidisciplinary-Social Sciences & Humanities (SSH). The existence of two areas called "Multidisciplinary" is due to the fact that in the OLIVA database there are a significant number of journals with a high degree of diversification and extent in their scope, crossing broader disciplinary boundaries. Thus, in group 7 we included journals that belong to at least two subject areas combining "hard" sciences with SSH. In group 8, we included journals that combine 5

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.
subject areas only within the Social Sciences and Humanities. This primary classification allowed us to reduce inconsistencies and to more accurately assess the disciplinary distribution of the journals analyzed.

Information on the publishers of each journal as well as the APC information was also taken from each journal’s website.

2. The scope of the OLIVA corpus and the institutional attachment of the 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, representing 87% of the records, although this percentage varies amongst the subject areas. There is an average of 3.1 authors per document and 3.3 per article. Given that multiple authorship is a more widespread practice in the Exact and Natural Sciences, Engineering and Biological Sciences, this data is already indicating that we are dealing with a production that has a broad disciplinary profile. One might think that the journals published 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. But we will see later that this is a multilingual corpus.

The 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 production is reflected even more prominently in the number of authors, given the growing trend in that country towards collaboration among 3 or more researchers, including in the social sciences and humanities. In terms of number of articles, next after the journals from Brazil are those from Mexico, Colombia, Chile, Cuba and Argentina. The journals from the rest of the region represent only 8.7% of the articles.

Table 1. Basic data on the OLIVA corpus. Number of journals, documents, articles and author records by country.

<table>
<thead>
<tr>
<th>Country of publication</th>
<th>Journals</th>
<th>Documents</th>
<th>Document authors</th>
<th>Articles</th>
<th>Article authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil</td>
<td>506</td>
<td>444,332</td>
<td>1,579,011</td>
<td>396,293</td>
<td>1,476,492</td>
</tr>
<tr>
<td>Colombia</td>
<td>291</td>
<td>102,762</td>
<td>241,335</td>
<td>90,530</td>
<td>224,630</td>
</tr>
<tr>
<td>Mexico</td>
<td>283</td>
<td>120,475</td>
<td>299,471</td>
<td>100,355</td>
<td>273,002</td>
</tr>
<tr>
<td>Argentina</td>
<td>167</td>
<td>46,237</td>
<td>121,052</td>
<td>35,919</td>
<td>104,093</td>
</tr>
<tr>
<td>Chile</td>
<td>144</td>
<td>69,095</td>
<td>199,308</td>
<td>57,032</td>
<td>178,521</td>
</tr>
<tr>
<td>Venezuela</td>
<td>97</td>
<td>34,939</td>
<td>92,097</td>
<td>30,161</td>
<td>85,814</td>
</tr>
<tr>
<td>Cuba</td>
<td>82</td>
<td>46,052</td>
<td>160,371</td>
<td>41,621</td>
<td>149,736</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>48</td>
<td>16,313</td>
<td>35,322</td>
<td>14,816</td>
<td>33,199</td>
</tr>
<tr>
<td>Peru</td>
<td>37</td>
<td>13,902</td>
<td>40,558</td>
<td>11,773</td>
<td>37,066</td>
</tr>
<tr>
<td>Uruguay</td>
<td>25</td>
<td>4,756</td>
<td>14,914</td>
<td>3,680</td>
<td>12,866</td>
</tr>
<tr>
<td>Bolivia</td>
<td>22</td>
<td>4,491</td>
<td>10,080</td>
<td>3,629</td>
<td>8,767</td>
</tr>
<tr>
<td>Ecuador</td>
<td>11</td>
<td>3,440</td>
<td>5,631</td>
<td>2,877</td>
<td>5,002</td>
</tr>
<tr>
<td>Puerto Rico</td>
<td>5</td>
<td>1,253</td>
<td>1,841</td>
<td>816</td>
<td>1,382</td>
</tr>
</tbody>
</table>
Concerning the temporal coverage of the available data, the corpus reveals more than a century of indexed Latin American scientific production. 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 much fewer documents in 2019 (0.67% of the articles) due to the fact that the data collection took place between May and June of that year, thus representing only a partial year. The earliest documents are from the *Memórias do Instituto Oswaldo Cruz*, which has 582 records between 1909 and 1939. Then we find the first documents of the *Revista Chilena de Pediatría* appearing in 1940, *Bragantia* in 1941, *Arquivos de Neuro-Psiiquiatria* in 1943 and *Anais da Escola Superior de Agricultura Luiz de Queiroz*, today named *Scientia Agriculta*, in 1944. The first appearances of journal documents from the other 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 Social Sciences and Humanities with a record in the database is the *Revista de Administração de Empresas* (1961) followed by the *Revista Latinoamericana de Psicologia* (1969).

Up to 1942 there are less than 100 documents per year; up to 1990, less than 1,000; and up to 1999, less than 10,000. The highest number is reached in 2015, with 58,025 articles. These values should be interpreted with caution because they are not necessarily representative of the history of journals in Latin America and the Caribbean. 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 the year of the indexing of a journal by that time seems exceptional.

**Table 1**

<table>
<thead>
<tr>
<th>Country</th>
<th>Journals</th>
<th>Documents</th>
<th>Authors</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panama</td>
<td>1</td>
<td>408</td>
<td>427</td>
<td>333</td>
</tr>
<tr>
<td>Dominican Republic</td>
<td>1</td>
<td>527</td>
<td>877</td>
<td>469</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>1,720</td>
<td>908,982</td>
<td>2,802,295</td>
<td>790,304</td>
</tr>
</tbody>
</table>

Source: The OLIVA database as of June 2019. Note: the column “Journals” includes active and inactive journals that were indexed by the time the documents were published; the column “Documents” includes articles and other types of literature; the columns for “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.

---

**Figure 1. The OLIVA Corpus (1995-2019): journals with at least one article published per year; total articles published per year.**
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 in the Social Sciences. As seen in Figure 1, the peak is reached in 2015 which 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 incorporation of the new issues into the Redalyc and SciELO databases. This delay varies widely from one publisher to the next and among countries, and is due to a multitude of reasons such as delays in publication or in the management of the databases, difficulties in XML-JATS markup, or differing priorities which the publishing institutions place on the inclusion of new issues in each database.

Accordingly, the evolution of the journals shows two significant increases, one in 1998-2000 and another in 2005, the years when SciELO and Redalyc were launched. The 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 activity also show a similar trend, although less pronounced than that of Brazil (Figure 2). The first big increase occurs two years later, between 1999 and 2000 when SciELO began to more actively promote itself outside of its country of origin (Brazil). For example, for Mexico there are 23 journals in 1996 and 87 in 2000. The most pronounced increase appears to be for the journals edited in Colombia when, in the same period, the number of journals goes from 4 to 35.

The impact of the appearance of Redalyc in 2005 can be noted for all countries, especially for Mexico and Colombia. The latter showed a very sustained growth in the number of journals in the following decade, eventually surpassing Mexico which experienced a slight stagnation in the number of journals. Argentina, Venezuela, and Cuba follow a similar
pattern with two sudden expansions. However, the increase in subsequent years is moderate compared to the other cases. More specifically, for journals published in Cuba, there is an upward trend between 2005 and 2015, followed by a decline. A decrease in the number of journals indexed in Venezuela can be observed beginning in 2009 and, more rapidly, since 2014. Finally, the case of Chile is also distinct with a sustained tendency, where 2005 did not imply a sudden change in the growth trend of the journals.

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

![Graph showing journals by country of publication from 1995 to 2017](image)

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 the journals that make up the OLIVA database. Journals from the Social Sciences 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 country of publication, 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 have their greatest relative presence in Argentina, Bolivia and Brazil. Figure 3 shows the disciplinary areas of the journals and allows for comparing their distribution in less visible countries such as those in Central America and the Caribbean with the large players in the region.
The comparison by country is very interesting but it is also relevant to observe publication asymmetries within the countries. Scientific production is often centralized and concentrated in the principal cities, in particular, in the biggest institutions of higher education. For this reason, we are interested here 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 great number of the journals: 38% of the total are published by only 53 mega universities, while the remaining 62% are 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.

<table>
<thead>
<tr>
<th>Publishing Institution</th>
<th>Number of journals published</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universidad Nacional Autónoma de México</td>
<td>61</td>
</tr>
</tbody>
</table>
As can be seen, the majority of journals 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 carry considerable weight. An exceptional case is that of Colombia, where publishing 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 in terms of number of journals within the OLIVA database.

3. Journal management and financing: the incidence of the APC

Let us now look at the distribution of the publishing institutions according to their editorial management and financing models. We found five types of publishers in 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 bring together small and medium size national and local specialized publishers), and finally; 5) the big publishers. The latter refers to Reed-Elsevier, Wiley-Blackwell, Springer and Taylor & Francis, considered as oligopolistic commercial publishers since they concentrated between 50% and 70% of the world’s scientific production up to 2013 (Larivière, Haustein y Mongeon 2015).

---

ECIMED was founded in Cuba in 1998 as the publishing house of the Centro Nacional de Ciencias Médicas which by then was already editing 15 journals. The center was part of the development of the medical sciences promoted by the Cuban government. ECIMED also publishes books and its official website functions as an open access portal for the 38 journals available in full-text.
As can be seen in Figure 4, 66% of the journals are edited by universities, close to 7% by governmental agencies, while scientific societies and independent research centers account for 20% of the publishers. 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 the share by ECIMED is noteworthy.

Most of the journals managed by the major commercial publishers are edited in Brazil (29), 7 in Chile, 3 in Colombia, 2 in Mexico, and 1 in Argentina. Described by discipline, this group (43) includes 20 journals that belong to Medical and Health Sciences, 14 to Exact and Natural Sciences, 5 to Engineering, 1 to Agricultural Sciences, and 3 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 the publication fee information from the websites of each of the journals. In most cases, the journals did not report publication fees. Upon further analysis, we verified that the absence of this information was due to the fact that they were indeed open access diamond journals, i.e. they do not charge authors for publishing or readers for accessing the contents. In our universe of 1,720 journals, we observed that 10% of the journals charge APC’s, the majority being published in Brazil (60% of the total 182 journals with APC). This is different from what was observed by Appel and Albagli (2019) in considering a larger universe of journals edited in Brazil, where the incidence of APC was still very low. In the OLIVA database, Brazil appears as leading the 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 which represent only 11% of the total of journals with APC’s (more than half of these belong to Psychology). No journal with APC was detected in the Humanities, that is why this discipline does not appear in the figure 5.

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

Amongst the journals with APC’s, 38% of them are managed by scientific societies, 37% by universities, 13% by the major commercial publishers, 7% by a governmental organization, and 5% by small commercial publishers. 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 of the sustainability of the journals and their varying degrees of institutionalization. Córdoba González (2021) observes that Latin American journals with APC’s have, by and large, low fees. This is also noticed in the journals included in the OLIVA database, which suggests that they are responding to the need to cover the costs of editorial processes and the management of the peer review process that are not funded by the institutions. The journals indexed in SciELO and Redalyc have specific cataloging, evaluation and digitization procedures which require material and human resources. These are generally provided by the publishing institutions, but in some cases they turn to the APC to pay for the DOI, translations, or administrative or technical support. This cannot, however, be equated to the business model of the commercial journals. Furthermore, journals with high APC’s, similar to the levels found in the mainstream industry, are exceptional in the region.
The sustainability of diamond access journals is a contested issue. The recently published OPERAS report highlights the importance of this model 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 that we have been describing – the vast majority of which do not charge for reading or publishing – demonstrate just the opposite: that they are not merely volunteer enterprises. In Latin America there is considerable consensus 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 not lie in the adoption of 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).

4. Trends in collaboration and language of publication.

Collaboration is an interesting and complex phenomenon in the production published in the journals included in OLIVA. First, it should be noted that only 25.8% of the articles are authored by a single author. This varies considerably among the journals of the different disciplinary areas (see Table 3). It represents 77.2% of the articles in the Humanities and 49.3% in the Social Sciences. In the Humanities, only 19% of articles are authored by 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% with four to eight. Table 3 shows that Multidisciplinary SSH journals have an intermediate percentage of authors compared to the two separate areas.

These trends within the social sciences undergo some variations if the analysis is segmented by the country of publication of the journals. Journals edited in Brazil, Cuba and Puerto Rico have, on average, 38.6% of their articles with only one author, a value significantly lower than that of the area as a whole. At the other end of the scale, the journals in the Social Sciences of Bolivia and Ecuador are largely dominated by 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 can also be noted but to a lesser extent. In Cuba, 64.2% of articles in journals in the Humanities have only one author. In Brazil, the percentage is 70%, which is not far from 70.1% in Uruguay or 73.6% in Chile, for example. In Argentina it is more than 80%, while Mexico reaches 87.6% and Costa Rica 89.2%.

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

<table>
<thead>
<tr>
<th>Disciplinary Area</th>
<th>One author</th>
<th>2-3 authors</th>
<th>4-8 authors</th>
<th>More than 8 authors</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agricultural Sciences</td>
<td>3.4%</td>
<td>27.6%</td>
<td>65.9%</td>
<td>3.1%</td>
<td>100%</td>
</tr>
<tr>
<td>Medical and Health Sciences</td>
<td>8.7%</td>
<td>30.7%</td>
<td>55.9%</td>
<td>4.7%</td>
<td>100%</td>
</tr>
<tr>
<td>Natural and Exact Sciences</td>
<td>11.5%</td>
<td>45.4%</td>
<td>40.7%</td>
<td>2.4%</td>
<td>100%</td>
</tr>
<tr>
<td>Engineering and Technology</td>
<td>8.0%</td>
<td>52.0%</td>
<td>39.2%</td>
<td>0.8%</td>
<td>100%</td>
</tr>
<tr>
<td>Multidisciplinary</td>
<td>17.7%</td>
<td>39.0%</td>
<td>40.2%</td>
<td>3.1%</td>
<td>100%</td>
</tr>
<tr>
<td>Social Sciences</td>
<td>49.3%</td>
<td>40.8%</td>
<td>9.7%</td>
<td>0.2%</td>
<td>100%</td>
</tr>
<tr>
<td>Humanities</td>
<td>77.2%</td>
<td>19.0%</td>
<td>3.7%</td>
<td>0.1%</td>
<td>100%</td>
</tr>
</tbody>
</table>
In general terms, the articles in the journals in the Agricultural Sciences, Medical and Health Sciences, and Multidisciplinary have mainly 4-8 authors. Articles with two or three authors are found primarily in journals in the Exact and Natural Sciences, and in Engineering and Technology. In Multidisciplinary, they represent approximately 40%.

Some countries stand out because their journals follow a different trend. In Brazil, in the Medical and Health Sciences journals, articles with only one author account for 5% of the articles, while those with 4-8 authors represent 61.8% of the total. In all other areas, the journals of Brazil show the same tendency of having articles with a greater number of authors. Another interesting case is Colombia, with a pronounced concentration of articles with two or three authors (instead of 4-8 authors as the rest). In Agricultural Sciences it totals 56.3% of the 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.

The degree of international collaboration can be observed through co-authorship. Figure 6 shows, for six countries, the proportion of articles authored 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 authored 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 with some type of international collaboration (9.8%). On the other hand, articles with only one Brazilian author are relatively few (14.1%). Here, collaboration amongst researchers affiliated to Brazil dominates (79.6%) and is distributed throughout the geography of the largest country in South America (see next section). It is interesting to note that international collaboration is a very limited practice for Brazilian authors when they publish in journals edited in Brazil. The focus on national collaboration is hence combined with a lower propensity for single authorship, as already mentioned.

Figure 6. Articles with country affiliations in 6 selected countries, by type of authorship and type of collaboration. (n=492,308).
It is interesting to compare these forms of collaboration that we have been observing in OLIVA with other databases where international collaboration of authors from these countries is much higher, and trending upwards. According to data from SCOPUS, 23.9% of articles authored by at least one author of Brazil in 2011 showed some form of international collaboration. This increased to 29.6% in 2015, and in 2020 reached 36.1%. In the case of Argentina, the values for the same years are 40.4%, 42.8% and 49.9%, respectively. However, it is worth remembering that Brazil’s share is significantly higher than that of the rest of the countries in the region (according to Scopus, articles by authors of Brazil represent 48% of the documents produced in Latin America in the period 1996-2020).

It seems likely that the degree of international collaboration by Brazilian researchers in these mainstream journals is affected by this country’s linguistic isolation within Latin America and its growing tendency to publish in English, as well as its internationalization policies prioritizing the North America and Europe. Meanwhile, for Spanish-speaking Latin American authors, there is a more sustained tradition of intra-regional collaboration. Based on the OLIVA corpus, it can be said that Brazilian authors collaborate more with each other, while 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 proportion as co-authors of their colleagues in journals edited in other countries. For example, of the segment of articles with authors from Argentina with international collaboration (n=3,906), there is also at least one author from Brazil in 26.5% of the cases. In collaborative papers with authors from Chile (n=5,376), co-authorship with Brazil is found in 19% of the articles. In the case of Colombia (n=6,729) co-authorship with Brazil reaches 16.6% of the articles. Clearly, national collaboration for Brazilians appears dominant when analyzed for the journals edited in Brazil, yet they have a significant participation in the co-authorship of articles published in other countries. We will further explore the mapping of international collaboration.

---

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 to a greater extent than with authors from Spanish-speaking countries depending on the recognition of the journal.

In the case of authors from Chile, it is worth noting some nuances in the general trends in the OLIVA database. The proportions of articles with individual authorship or with international collaboration are more than double than those of the journals of Brazil, with a lower incidence of collaboration within the country itself. Co-authorship with colleagues located in other countries is the highest of all, a feature which, as explained below, may be indicative of an academic culture with peculiarities within the region. Among the remaining cases, Cuba presents similarities with Brazil although with a greater degree of international collaboration. On the other hand, Colombia and Mexico are similar to Chile, although with a greater proportion of national collaboration.

Along similar lines, trends analogous to those mentioned above can be analyzed by observing articles by authors of different countries and the country of the journal (see Figure 7). The first thing that emerges from this analysis is that, for the countries in question, articles are published above all in journals of the same countries as the authors. This feature is particularly pronounced in the case of authors from Brazil: 93.7% of the articles which they author are published in journals of Brazil. Chile comes second, yet it represents only 1.6% of the total of articles. In bibliometric studies based on the mainstream indexing sources, this characteristic is generally interpreted as academic inbreeding, thus frequently resulting in a negative assessment. In the following section we propose an alternative approach which contrasts with simply attributing this national collaboration to practices of academic inbreeding. It is based on a detailed study of the relevant number of articles authored by Brazilians and published in Brazil.

The case of the authors from Chile, 75% of whom have published in journals of their own country, is certainly a high value but lower than other cases. Journals from Colombia, Brazil and Mexico are important recipients of Chilean authors’ submissions within the region. This type of distribution may be linked to greater pressure from the assessment culture of Chilean institutions which discourages publishing in the nationally edited journals and provides direct incentives for international collaboration, as was already mentioned in the discussion of Figure 3. On the other hand, authors from Colombia, Cuba and Mexico tend to make their submissions to their nationally edited journals, although not at the same degree as described for Brazil.

Argentina also has its own peculiarities. The fact that it has the lowest share (56.4%) of articles published in its journals with at least one author from Argentina suggests a stronger trend to publish abroad than that observed for the Chilean case. However, there are other explanations. There is evidence that journals from Argentina are underrepresented in the OLIVA database, a product of a traditional selection of social science and humanities journals edited in that country by the services of the Latindex catalog. This selection was driven, among other factors, by the characteristics of the existing assessment culture at CONICET (the National Science Council) and the national universities where publications indexed in Latindex are rewarded at the same level as those included in SciELO, Scopus and WoS (Beigel, 2014). To this we can add the lower technical complexity of indexing in the Latindex Catalog 2.0 compared to the
requirements of SciELO and Redalyc, and the clout of the national (Argentine) node of Latindex. If we compare the journals of Brazil, Argentina, Chile, Mexico and Colombia in the Latindex Catalog 2.0 with the OLIVA database, we see that in the case of Chile there are 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, that is, almost twice as many. In contrast, the journals of Argentina in the OLIVA database that do not belong 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.

Let’s move on now to language trends. In contrast to the highly predominant role of English widely seen in mainstream databases, the OLIVA corpus is characterized by a remarkable linguistic diversity, driven by the infrastructure developed in the region with multilingual indexing and publishing protocols and infrastructures. Table 4 shows the central role of Spanish and Portuguese, but also a significant presence of English which represents 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 journals of Brazil were published in Portuguese, while English represents one-third (34.6%). This value is only comparable to the percentage of English in articles in journals of Puerto Rico (50.5%), Chile (23.7%) and Mexico (17.1%). In the case of Chile, it might be one more
element pointing to an academic culture more influenced by mainstream standards. In the case of Mexico and Puerto Rico, it may be influenced by traditional academic ties.

Outside of the journals of Brazil, Mexico, Puerto Rico and Chile, an average of 5.4% of the articles are in English. On the other hand, Portuguese as a language of publication outside of the journals of Brazil journals is very low, reaching 5.4% of the articles in the journals of Costa Rica, 3.8% in journals edited in Uruguay, and 2.5% in Ecuador. Conversely, in the Spanish-speaking area, Spanish is by far the predominant language. In journal articles from Bolivia, Cuba, Ecuador, Peru, Dominican Republic, Uruguay and Venezuela, Spanish accounts for more than 90%.

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

<table>
<thead>
<tr>
<th>Language</th>
<th>Nº of Articles (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish</td>
<td>43.7%</td>
</tr>
<tr>
<td>Portuguese</td>
<td>32.09%</td>
</tr>
<tr>
<td>English</td>
<td>23.91%</td>
</tr>
<tr>
<td>French</td>
<td>0.2%</td>
</tr>
<tr>
<td>No data</td>
<td>0.12%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>

*Source: The OLIVA database (June 2019).*

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

In journals from countries where Spanish is the official language, the share of English among articles in journals of the Natural and Exact Sciences is 41.9% and for Engineering and Technology, 28.4%. Portuguese is very much in the minority, and only in the Social Sciences does it 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%).

Publication in English is associated with a greater number of authors. The average number of authors in the English-language articles is 4.3. For Portuguese it is 3.3 and for Spanish, 2.7 authors per article. The relatively high number for Brazil is also linked with the tendency of authors and journals of Brazil to publish articles with two or more authors. The same decreasing order (English, Portuguese, and Spanish) in the average number of authors is observed in all areas.

There is also a relation between the authors’ country affiliation and publication 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 identified) 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 types of
The evolution over time of the distribution of languages of publication can also be analyzed. 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 linguistic contexts - Portuguese and Spanish - reveals some trends that are worth highlighting. In the case of journals from Brazil, the share of articles in English went from representing 24% in 2009 to 58% in 2018. The decrease in the relative weight of articles in Portuguese over this period was of the same magnitude, while Spanish remained at a modest 3%. This trend of English is repeated in all disciplinary areas, but very attenuated in the Social Sciences and Humanities, while it is strongly endorsed by the Agricultural Sciences, Medical and Health Sciences, Engineering and Technology, and Exact and Natural Sciences. The predominance of English was already noticeable at the beginning of the period as represented in the journals of the latter disciplinary area.

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 represented only 0.3% of the articles.

In the other countries, although the evolution over 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 segmentation by discipline. In the journals in Agricultural Sciences, Engineering and Technologies and Natural and Exact Sciences considered as a whole, English went from representing 29% of the articles in 2009 to 42% in 2018. At the same time, Spanish dropped from 70% to 57% of the articles. In other words, the official language of most Latin American countries remained the predominant language of the articles.
The other disciplinary areas, on the other hand, also show similar behavior. Overall, articles in Spanish in these journals went 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 marked in the cases of journals from Uruguay and Colombia.

5. National collaboration in Brazil.

As previously mentioned, the weight of Brazil is central to the analysis of the scientific journal space in Latin America. Journals of Brazil account for 29.4% of the 1,720 journals included in the OLIVA corpus. Looking at the journals by discipline, most are found in 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 Technology (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 the articles in the OLIVA database have been published in journals of Brazil. 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%).

In this last section, a complementary data source is analyzed for an in-depth characterization of the national collaboration in this country. This is the metadata of articles published in the journals of the SciELO Brazil collection over the five-year period 2016-2020, totaling 104,750 documents. This information was extracted from SciELO in May 2021. This is the portion of the database that has undergone the process of standardizing the names of authors and their institutional affiliations, and allows for a more precise analysis. This set of documents includes authors from 162 countries. In 74.5% of the cases, there are only authors affiliated with Brazilian institutions (and 23.6% with only one author), while in 18.3% of the cases there are only authors from other countries. In just 7.1% of the documents is there collaboration between authors inside and outside of Brazil. The picture is therefore very similar to the one presented in Figure 6. Notably, in these articles co-authored by Brazilians and foreigners, only 17% are co-authored with Latin American countries. This phenomenon, to a large extent determined by linguistic aspects, is something that we have already commented on.

However, looking in detail at the authors from Brazil, a heterogeneous picture emerges. There are articles by researchers from all 27 Brazilian states, but strongly concentrated in only some of them. In the set of documents that have at least one author from Brazil with his/her geographic state declared (n=85,528), 85.7% belong to five states. These are the states of 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 the South of Brazil.

Of this segment of 58,148 documents, 63% are coauthored, and among these one can see some trends in within-country collaboration that opens up an interpretation beyond the dualistic terms of academic inbreeding/cross-breeding. In effect, 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, making it difficult to characterize a collaboration between two Brazilian authors (or publication in a journal of their country) as a practice of academic inbreeding. Collaboration between authors from different institutions within the same state can be thought of in the same vein.

In the absence of data that could illustrate inter-institutional collaboration, the analysis of inter-state collaboration in the case of Brazil introduces a significant nuance in 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 only by authors from Brazil, and that other countries show lower percentages of publication with authors only from within their own country, such as Argentina (48.3%), Chile (52.7%) and Mexico (56.5%). The documents that we analyze in this section (SciELO 2016-2020) show that, when there is participation of authors from Brazil, in 41% of the cases there is collaboration, either by authors from two states (31.9%) or from three or more states (9.1%). The remaining 59% (only within-state collaboration) is presented as a proportion similar to that of exclusively within-country collaboration in the Spanish-speaking countries mentioned.

In effect, 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 the Southeastern states in intra- and inter-regional collaborations within Brazil. Although such states continue to have this central role - particularly São Paulo - in recent years collaboration between other states in the same or different regions has increased, a phenomenon that is particularly strong in the Northeast and the South.

If one were to look at the research system of the magnitude of Brazil at its highest level of aggregation, one would lose the detail necessary to analyze the complexity of the interactions inside its boundaries. 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 present a greater degree of inter-state 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 the 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 instead between different states.

Table 5. Documents in SciELO-Brazil with two or more Brazilian authors, by inter-state collaboration and discipline of the journal, 2016-2020 (N=58.148).

<table>
<thead>
<tr>
<th>Disciplinary area</th>
<th>1 state</th>
<th>2 states</th>
<th>3 or more states</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agricultural Sciences</td>
<td>48.3%</td>
<td>37.6%</td>
<td>14.1%</td>
<td>100%</td>
</tr>
<tr>
<td>Natural and Exact Sciences</td>
<td>50.2%</td>
<td>37.3%</td>
<td>12.5%</td>
<td>100%</td>
</tr>
<tr>
<td>Biological Sciences</td>
<td>51.6%</td>
<td>35.3%</td>
<td>13.1%</td>
<td>100%</td>
</tr>
<tr>
<td>Multidisciplinary</td>
<td>55.0%</td>
<td>33.3%</td>
<td>11.7%</td>
<td>100%</td>
</tr>
</tbody>
</table>
Nonetheless, there are clearly important asymmetries in the different variants of collaboration between authors from different states. For authors of documents from a given state, the other top five states with which there is co-authorship are analyzed, São Paulo appears as a prominent “partner” in 26 cases, Minas Gerais in 25, Rio de Janeiro in 21, Rio Grande do Sul in 14 and Paraná in 12. Then Pernambuco comes in with 7 and Brasília with 5. There are seven states which 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 region. Table 6 shows the nine states of the region, and the types of collaboration of authors from each of the states with authors from the same or other states and regions in Brazil. First of all, there is a considerable tendency (just over 40%) to collaborate with other authors from the same state. However, the collaboration amongst 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 split approximately equally. The total co-authorship with authors in the remaining thirteen states ranges between 7.7% and 12.7%. In other words, the collaboration “outside state borders”/between states takes place with the same degree of intensity between states in the same region as it does between these same states and the five major centers of scientific production. Of course, this trend is complemented by the modest role of the other states in within-state collaboration.

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

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

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 observed variations in the table are not pronounced, which suggests that these are relatively stable features of the Brazilian scientific production space. 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. The main feature here is the degree of in-state collaboration which is more pronounced with respect to the case of the Northeast since in all cases it exceeds 50%. For each state in the table, the other four states represent approximately the same proportion of collaboration (fourth column as all the remaining states in Brazil (fifth column). Clearly, the concentration in collaboration in the institutions of these five states stands out. But it is also clear that, even though it is smaller, their collaboration with states less prominent in the Brazilian scientific sphere is not totally marginal. Finally, it is worth noting the remarkable similarity in the profiles of the five states in Table 7, which might be suggesting – as mentioned before – consistent trends in this field of collaboration.

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.

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


This look at 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 characteristic feature of the Brazilian scientific production space as analyzed through the SciELO documents of the last five years, that is, the period 2016-2020. This is a question of a practice of collaboration between teams located in clearly different spaces/environments, with a diverse academic history and processes of institutionalization that affect, in particular, some disciplinary areas since, as might be expected, inter-state collaboration is less prevalent among those disciplinary areas that tend to have fewer authors in their articles.

The concentration of the authors responsible for these documents in the institutions of the Southeastern and Southern states does not, however, ignore the significant collaboration with colleagues from the rest of the country’s extensive geography. In addition, states such as São Paulo, Minas Gerais or Rio de Janeiro can hardly be considered as homogeneous within each other. They include a diversity of institutions – federal, state, private – among which presumably there is also collaboration in terms of their publications. And this gives further nuance to the idea of academic inbreeding in terms of authors’ affiliation within the same state given that, for example, collaboration between authors from Unicamp (Universidade Estadual de Campinas) with those of USP...
(Universidade de São Paulo), both in the same state, can hardly be classified as a provincial interaction.

5. Conclusions

Many experts warned as early as the 1980s and 1990s that the science produced in Latin America was not visible to the world because it was rarely cited in "international" databases, which is why Gibbs (1995) spoke of a "lost science of the third world". In fact, it was in those decades that an infrastructure for scholarly communication was strengthened in the region which sought to raise the value of research outputs produced in this part of the world with its uniqueness in 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 in a fragmented way because, despite the best of institutional efforts, interoperable systems still have not been achieved to allow for the scientific production of Latin America to be known, visible and valued world-wide. It is with this objective in mind that the OLIVA project was developed at the Universidad Nacional de Cuyo (Mendoza, Argentina) to enhance the value of the production assessed and indexed in articles in such a way as to have a growing presence and impact in academic assessment 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 that these open access publishing platforms have made and the extent to which there now is a Latin American production with high standards of quality whose value is still under-represented in current academic assessment models. Perhaps one of the main findings that emerges from the analysis is the longevity and diversity of the journals, some of which were created in the late nineteenth and early twentieth centuries. There is also a diverse representation of disciplinary areas, modes 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 those collaborations that represent a combination of efforts from different universities and especially from different states, thereby examining the geospatial dynamics of national scientific production. The analysis of Brazil, 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

---

8 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 nineteenth century, such as: Gaceta médica de México (1864), Gaceta Médica de Caracas (1893), Revista chilena de historia natural (1897), Revista de Ciencias Agrícolas (1903), Boletín de la Sociedad Geológica Mexicana (1904), Ingeniería, investigación y tecnología (1908), Memorias do Instituto Oswaldo Cruz (1909), Revista industrial y agrícola de Tucumán (1910).
researchers from different states throughout the country. We also characterized the states that, by their own importance, play an important role in the corpus and yet show significant collaboration with colleagues from states with less weight in the Brazilian scientific system.

An examination of the publishing institutions of the journals in the OLIVA corpus 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 production 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%. The SciELO and Redalyc journals, in contrast, have an eminently academic trait, being published almost entirely by universities, scientific societies and government agencies, offering a platform of editorial services anchored in the public domain. In the region, the creation of journals has historically been in the hands of the scholarly community, which continues to support them materially even though the evaluation systems encourage their academics to publish in mainstream journals.

Another intrinsic feature of scholarly communication in Latin America is the widespread adoption of Open Access which is fostered by the decision of research funding agencies, universities, scientific societies, and publishers to provide not-for-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 websites of the SciELO Network and Redalyc platform, as well as on the websites of each journal, and in thematic and institutional repositories. In addition to the production analyzed in these 1,720 journals, there are thousands of other journals in the region that show evidence of long-term operation that have been key to expanding open access in the region.

The study of the production published in SciELO and Redalyc has allowed us to discuss some suppositions that permeate the recognition of scientific knowledge, in other words, that of the prioritization of the journals commonly characterized as "mainstream" while discrediting the rest. OLIVA reveals two dimensions of this scientific communication space of increasing 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 opposing duality of the "mainstream or global circuit" - led by the most prestigious journals - and the "regional or local circuits", with limited circulation. This duality, promoted in the context of the commercialization of journals and reinforced as a "should be" by the assessment systems through the use of the journals’ Impact Factor, has undervalued non-commercial open access publications such as those we are concerned with in this study. This perspective has also failed to appreciate the forms of intra-national collaboration whose complexity we seek to show in this paper for the case of Brazil, and which are of particular relevance for research addressing local issues that are enhanced in particular in this type of collaboration. We also point out the linguistic particularities of Brazil which explain, to a large extent, 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 attachment, show that this is a system that promotes broad dissemination and multi-scale research programs, which can very effectively address the
current needs of scholarly communication in times of open science. This is not only based on Latin American capabilities in terms of professional scientific publishing, indexing and open access, but also on its experience in producing interoperable regional indicators as demonstrated in the joint publication developed in collaboration with the main indexing systems in the region, PKP and CLACSO (Alperin, Babini and Fischmann, 2014). More efforts in this direction to improve the visibility and comparability of Latin American production will enable the region to counteract a major trend in two and, until now, separate but increasingly intrinsically linked forces: commercial threats to the open science movement and the urgent need for a reform of the traditional forms of research assessment. It is our hope that the OLIVA project will stimulate more and more collaborative spaces to promote a socially relevant and participatory science.

References


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.