

SciELO Citation Index and Web of Science: Distinctions in the Visibility of Regional Science

Diana Lucio-Arias¹, Gabriel Velez-Cuartas² and Loet Leydesdorff³

¹ *dlucioarias@gmail.com*

Colombian Observatory of Science and Technology, Bogota (Colombia)

² *gabrielvelezcuartas@gmail.com*

Grupo de Investigación Redes y Actores Sociales, Departamento de Sociología, Facultad de Ciencias Sociales y Humanas, Universidad de Antioquia; Calle 70 No. 52-21, Medellín (Colombia)

³ *loet@leydesdorff.net*

University of Amsterdam, Amsterdam School of Communication Research, Amsterdam (The Netherlands)

Abstract

In this study we compare the visibility and performance of Latin American and Caribbean (LAC) Science in terms of its presence in the core collection indexes included in the *Web of Science* (WoS) —*Science Citation Index Expanded*, *Social Sciences Citation Index*, and *Arts & Humanities Citation Index*—and the *SciELO Citation Index* (SciELO CI)—which was recently integrated into the WoS platform. The purpose of this comparison is to provide some inputs to reconstruct the role of SciELO as a communication platform for science produced in Latin America and the Caribbean, and to provide some reflections on the potential impacts—in terms of a better understanding of the global scientific scenery—of the articulation of SciELO CI into WoS: Are there significant differences in the region's scientific results when studied from publications included in SciELO CI versus those included in the traditional core collection of the WoS? Are regional exercises, such as SciELO, successful in enhancing the visibility of regional scientific production?

Conference Topic

Journals, databases and electronic publications

Introduction

Although the participation of Latin American and Caribbean (LAC)-edited journals in WoS has increased over time, this growth is not comparable to the growth in the participation of scientific articles with at least one author affiliated to an institution in LAC. This increase in participation has been interpreted as a successful integration of LAC science into the world repertoires despite a persistent and notorious gap in the making of good scientific journals (Meneghini, Mugnaini & Packer, 2006). The difference in the nature and characteristics of the journals considered and included in each of the indices justifies our expectation of finding significant differences in the science produced in LAC and communicated through WoS or SciELO CI indexed journals: while the inclusion policy of WoS targets the top quality journals by discipline, the program SciELO has had an inclusive policy aimed at increasing visibility and circulation of LAC journals and their content.¹

¹ SciELO (Scientific Library on Line) was a program that was initiated in Brazil in 1997 with the purpose of offering a core of Brazilian scientific journals in an open access mode through internet. The program had a successful expansion in the region and now includes, in addition to Brazilian, journals from Chile, Cuba, Spain, Venezuela, Colombia, Argentina, Costa Rica, Mexico, Portugal, Peru, and Uruguay. It is important to note that the SciELO program transcends the SciELO citation index which is the subject of this study. Not all the scientific journals that belong to the SciELO collection and whose content has been made available through SciELO's program belong to SciELO's citation index.

Another difference in the origins of SciELO and WoS that might be helpful in explaining the differences in regional scientific communication is related to the disciplinary context of each of the indexes. A lot has been written about the “natural” or hard sciences origin of WoS, which derived from the Science Citation Index (Garfield, 1971), but was expanded to include a broader range of journals and then accompanied by the Social Science Citation Index and later on by the Arts & Humanities Citation Index. The three indexes have been operative since 1978. SciELO, on the other hand, resulted from cooperation of the Fundacao de Amparo a Pesquisa do Estado do Sao Paulo (FASPEP) and the Latin American and Caribbean Center for Health Sciences Information (Bireme) of Panamerican and World Health Organization (PHO/WHO).

We believe that SciELO’s contribution to global science relies on its impact in the circulation of LAC scientific production and therefore the visibility of this production. In the last 15 years, SciELO played an important role in the development of capabilities in LAC to produce world-class scientific results, particularly through the consolidation of a regional base of high-quality scientific journals. The financial requirements to maintain such an exercise updated, expanding and relevant (Aguillo, 2014), together with the potential of SciELO indexed journals to provide a representation of LAC science, might explain the interest behind the inclusion of the regional exercise in the Thomson Reuters owned databases.

The inclusion of SciELO into WoS has had a mixed reception in the LAC scientific community. In 2007, an alliance between Scopus and SciELO raised expectations of all SciELO information to be included in Scopus (Elsevier, 2007). The potential impacts of the inclusion of the journals, and the ambiguity of whether all SciELO journals would be included in Scopus raised some concerns in the LAC scientific community. The negotiations behind SciELO’s inclusion either in Scopus or WoS, was perceived by some editors of LAC journals as a “sell-out” of SciELO’s principles and allowed uncertainty in the future of the regional journal structure that SciELO had aimed to consolidate.

With this paper we expect to contribute on the relevance of both indexes and the complementarities between them as they represent different styles of scientific communication that transcend the center-periphery debate on scientific production. This section is followed by a section in which we introduce the data and methods employed for this study. The results section will focus on the differences between the indices; specifically in the geographical, collaborative aspects, and cognitive characteristics of the communications in each. We finish this contribution with some reflections on the challenges and opportunities of the integration of SciELO into WoS.

Data and Methods

We downloaded all the bibliographical information from the core collection of the WoS (SCI expanded, SSCI, A&HCI) for 79,924 documents that responded to the search query for affiliation country to any LAC countries AND publication year 2012. The same information was downloaded for 30,518 documents that responded to the same search query in the SciELO CI available through WoS. While participation of LAC authors explains 73% of the total publications in SciELO CI, in WoS, this participation is lower than 5%.² The organization of the information into relational databases was possible through dedicated routines available at <http://www.leydesdorff.net/scielo> and <http://www.leydesdorff.net/software/isi/index.htm>.

²In January 2015, a total of 1,899,805 documents were included in WoS with publication year 2012, and 41,621 in SciELO CI.

In order to assess some of the differences in the sets of data considered in this analysis, we provide some descriptive statistics in Table 1. We include the mean and the standard deviation to provide some order of magnitude and dispersion among attributes.

From Table 1, differences among the types of communications included in each set are evident. The mean (μ), represents the average number of authors, addresses, citations, cited references and subject categories per document and the standard deviation (σ) is included to illustrate dispersion in these data. The documents in journals indexed in WoS have more citations, and more frequently result from collaborations among larger number of authors in European or American institutions. These documents are more codified (in terms of the cited references used) as well, and, in general, have a significantly larger impact (in terms of citations received). The mean and standard deviation of the journals are included to represent the average number of LAC documents per journal. Although fewer journals concentrate LAC scientific production in SciELO CI than that in WoS, dispersion among different titles is greater; as can be expected, SciELO CI indexed journals have a larger participation of LAC authors compared with authors from other countries. A total of 163 journals are indexed in both WoS and SciELO CI.

Table 1. Differences in the sets of LAC publications from SciELO CI and WoS Core collection.

LAC publications	SciELO CI			WoS Core Collection		
Records	30,518			79,924		
Statistics	N	μ	σ	N	μ	σ
Authors	91,269	3.8	2.4	306,560	14	144.3
Addresses	11,858	2.3	1.5	168,390	3.9	14.3
Times cited	7,733	0.3	0.7	274,225	3.4	18.6
Cited references	681,151	26.2	19.1	1,969,653	37	29
Subject Categories	186	1.2	0.7	246	1.5	0.8
Journals	750	40.7	44.5	7,268	10.9	28.0

We use the Overlay maps Toolkit available at <http://www.leydesdorff.net/overlaytoolkit> (Rafols, Porter & Leydesdorff, 2012) to provide the different visualizations of the relations among disciplines in each of the document sets (SciELO CI and WoS core collection). We rely on these visualizations to suggest disciplinary differences in each of the sets of documents. We expect some of these differences to reflect on diverse goals and interests in the management of each of the indices and which were shortly introduced above.

To reflect upon the distinctions in the collaborative nature of the communications in each index, we build a collaboration network between countries using Pajek.

Results

In this section we provide some results on the differences between communications in the Core Collection of WoS and the recently integrated SciELO CI, focusing on the regional, collaborative and cognitive aspects underlying these communications. In Table 2, we provide the number of records in each of the sets by country of origin of the authors. To normalize for documents with a high number of co-authorships we include a fractional counting of documents considering the total number of signing authors.

The divergence in the countries' participation in the scientific production of LAC can result from (a) the degree in which the specific country has become articulated in the SciELO program and the efforts in increasing the SciELO journal list of each country. As can be expected, the most important SciELO journal collection is from Brazil and it includes 337 journal titles, Colombia follows with a total of 184 journal titles, Mexico has 149, Argentina

and Chile 107 and 106 journal titles each. Another explanation is (b) the specific country's treatment and importance of national scientific journals.

The policy effort supporting national scientific journals varies in the region where some countries privilege international publication while others aim at balancing international visibility with support to local journals and local publishers (Vessuri, Guédon & Cetto, 2014). Different publication strategies are also evident from Table 2 where the effect of fractional counting seems to be more drastic for communications in journals indexed in WoS Core collection than in SciELO CI. Colombia, for example, has relied on collaborating with international peers to increase their participation in international journals and databases (Lucio-Arias, 2013).

Table 2. Regional distribution of papers in WoS Core collection and SciELO CI.

Country	SciELO CI		WoS	
	Records	Fractional	Records	Fractional
Brazil	19,537	11,929.5	44,812	21,844.1
Colombia	3,065	2,312.2	4,007	1,734.9
Chile	2,409	1,754.3	7,277	3,562.0
Mexico	2,336	1,529.2	13,041	5,879.3
Cuba	1,979	1,053.5	966	320.8
Argentina	1,625	1,223.8	9,975	4,953.8
Venezuela	526	340.8	1,240	543.9
Peru	480	344.0	975	336.1
Costa Rica	284	189.4	514	310.8
Uruguay	99	51.8	868	195.3
Ecuador	53	25.0	465	153.4
Bolivia	42	20.0	85	17.0
Guatemala	23	11.4	52	8.0
Panama	22	8.0	416	120.7
Puerto Rico	22	8.0	N/A	N/A
Paraguay	27	10.7	43	6.1
El Salvador	11	5.1	24	3.1
Jamaica	10	3.1	9	1.8
Nicaragua	20	8.4	31	4.3
Honduras	3	1.0	25	2.8
Dominica	1	0.2	2	0.4
Dominican Republic	1	0.2	33	4.4

The alliances and collaborations reflect important differences in the networks of collaboration that emerge from LAC scientific communications in each of the indices considered (See Figures 1 and 2).

Collaborations in WoS suggest the importance of North America and Europe as allies in the production of scientific knowledge in the region. Collaboration of LAC countries with peers “from the north” dominates scientific communications where LAC participate. Regional collaboration seems not very relevant and in fact not as important as collaboration with Asia, Africa and Oceania. South-South collaboration has received a lot of attention (Arunachalam & Doss, 2000; Chandiwana & Ornbjerg, 2003) and has become an important issue in the

development policy agenda.³ We believe, nevertheless, that South-South collaboration depicted in Figure 1 is mostly mediated by developed countries and does not represent necessarily a transfer and exchange of resources and knowledge.

The resulting map of collaborations in LAC scientific communications in journals indexed in SciELO CI, suggest a more pronounced strategy based on the regional conjugation of research efforts. Collaboration with Europe is mainly oriented towards Spain and Portugal, suggesting language and cultural similarities as a strong motivation to collaborate. Collaboration with North America and particularly with the United States might rely on geographic proximity as this is stronger in the case of Mexico.

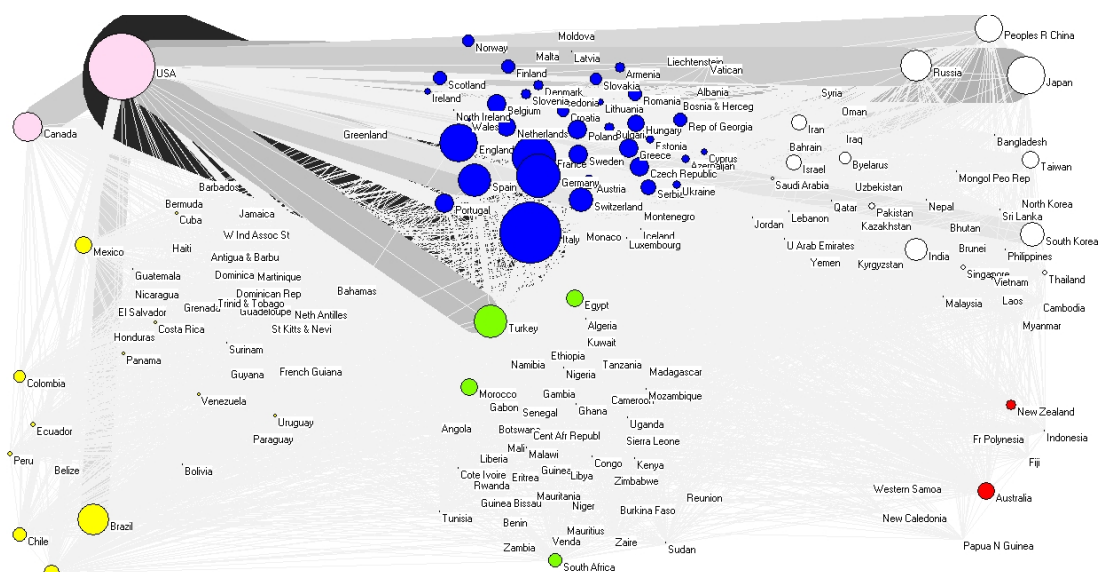


Figure 1. International Collaboration from LAC communications in WoS Core Collection.

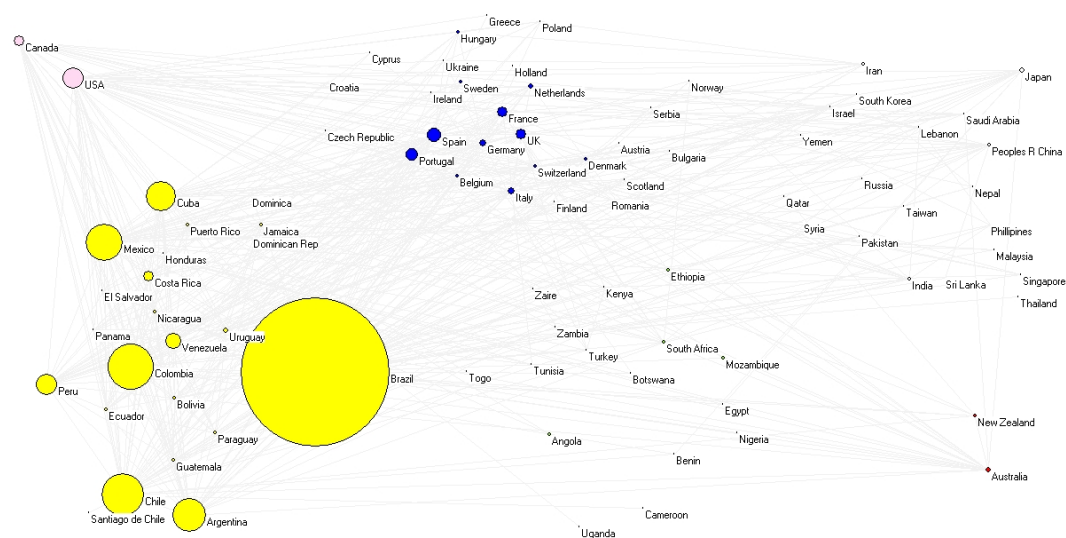


Figure 2. International Collaboration from LAC communications in SciELO CI.

Although it deserves further research, we expect collaborations in SciELO to be a better representation of South-South cooperation, which implies an exchange of resources and ideas within developing countries to solve similar development problems. Collaboration in Figure 2

³ There is a United Nations Office for South-South cooperation with a website at <http://ssc.undp.org/content/ssc.html>.

within LAC, Africa and Asia might be a better representation of South-South cooperation. We expect less mediation of the North in the South-South collaboration for the case of SciELO CI indexed communications.

In summary, the differences between Figures 1 and 2 suggest distinct communication practices when (a) aiming at results with international visibility than when the main goal is (b) regional or local diffusion of scientific results through regional journals. While for WoS (Figure 1) strong ties can be indicated with North America and Europe, regional collaboration seems dominant in Figure 2. The participation of the USA in Figure 1 and Brazil in Figure 2 should be interpreted considering that these countries have the highest numbers of indexed journals in each of the respective databases.

This can also result from the different disciplines represented in each index. While WoS has some dominance of “hard” sciences, which are more prone to be published in English and in collaboration, for SciELO CI the disciplinary participation seems to favor the social sciences (see Figure 3 and 4).

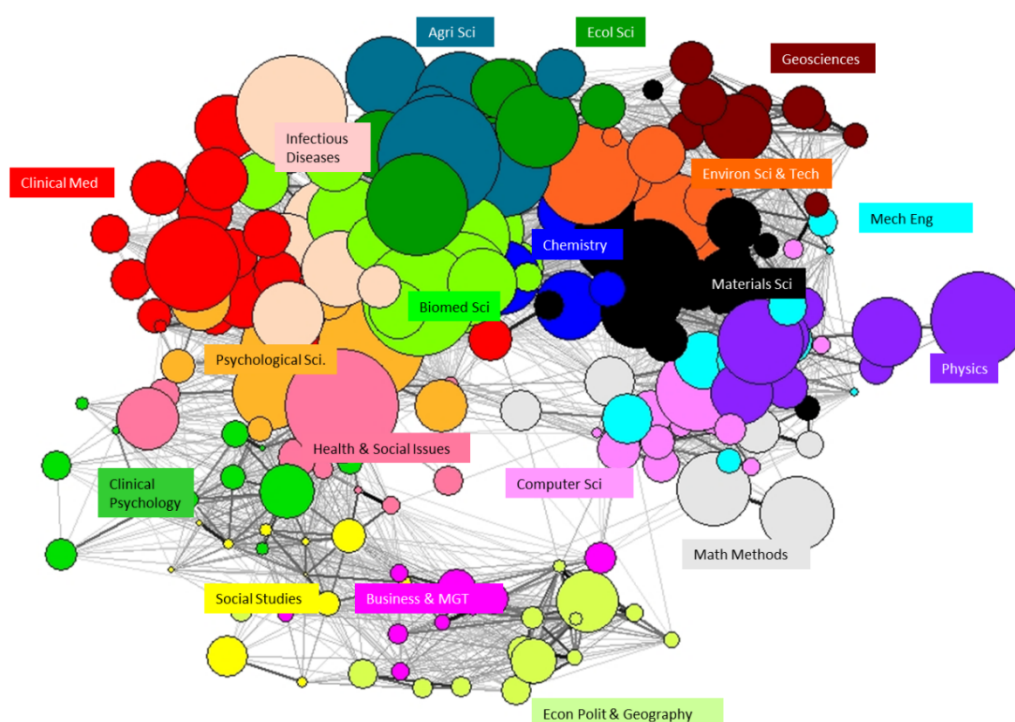


Figure 3. LAC map of Science, WoS Core Collection; 224 Web of Science Categories.

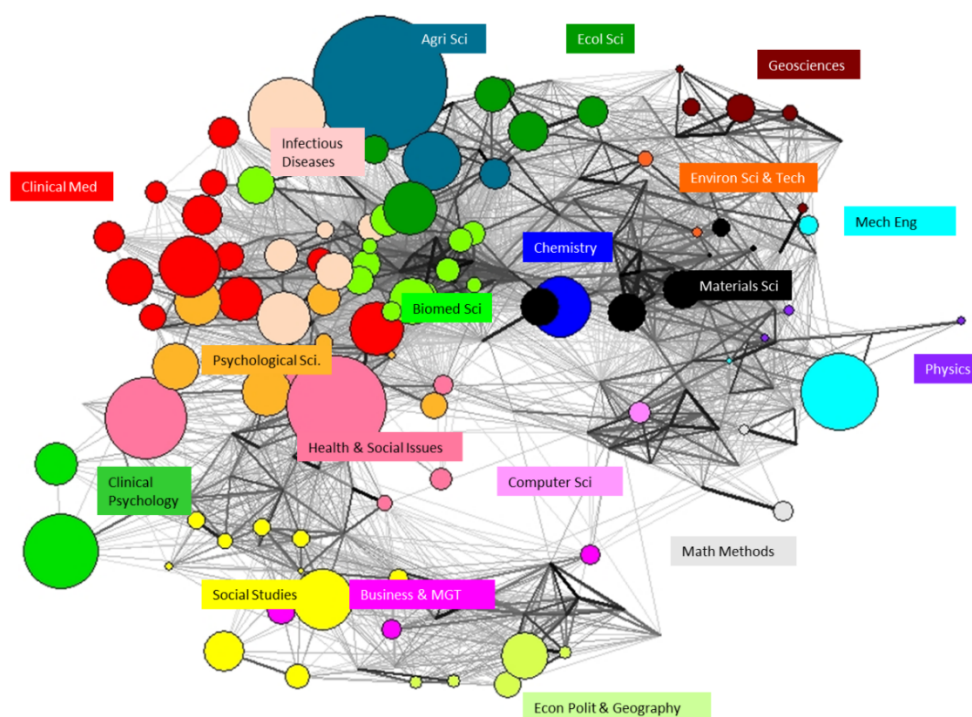


Figure 4. LAC map of Science, SciELO CI; 224 Web of Science Categories.

Figures 3 and 4 suggest differences in the thematic orientation of the communications in each index. Contributions from the natural sciences are better represented in WoS Core Collection; nevertheless, SciELO CI provides a valuable insight into the regional scientific production in the social and health sciences (where social aspects of the health and medical sciences like research in public health has a better representation), and agriculture. Our expectation is that in-depth analysis of the subjects addressed by the communications would exhibit differences in the sets; communications in SciELO CI will address topics of regional relevance.

Reflections and Further Work

In the last twenty years, scientific development together with technological change and productive innovation have raised interest in the LAC countries, and as a consequence been targeted on the public-policy agenda. Important aspects in the institutionalization of scientific research, such as the consolidation of public institutions for the promotion of science technology and innovation, strengthening of public research institutes, the growth of PhD programs, and the formation and formalization of a journal structure, to socialize scientific results obtained in the region, have also characterized these last decades.

Although growth in the participation of LAC scientific production in traditional databases, such as Web of Science and Scopus, has also been the norm in this period, a common concern in the community has been the challenges to properly socialize scientific results when they are of little interest for mainstream scientific journals. The perseverance in LAC scientific communications of Spanish and Portuguese, as the main languages for communication, particularly in sciences with an important social component, demands alternative means of communication outside international journals as they might have their own structures. Leydesdorff and Bornmann (in press), for example, found a specific citation pattern of Spanish and Portuguese journals in library and information sciences (LIS).

This demand has been acknowledged and as a consequence, most LAC countries have an important structure of national journals. This poses other types of challenges in terms of

research assessment and evaluation. While rankings of international journals and measures based on citations allow researchers and librarians to make informed decisions on the expected quality of a scientific journal's content, this distinction is more difficult and in occasions impossible when considering national publications. The proliferation of local journals edited by faculties or departments for the diffusion of mainly their own researchers' findings makes the distinction among journals harder.

The need to assess and monitor research results comes together with the demand for a transparent classification among scientific communications. How to assess scientific communications included in international journals versus regional or national journals? In part as a response to this need, different LAC countries have joined the SciELO program. SciELO, in our perspective, has had a positive impact on the consolidation of regional research capabilities and in providing a proper infrastructure for regional exchange and communication.

As was suggested in the collaboration networks analyzed, the SciELO program seems to have transcended the LAC region and includes authorships from Africa and Asia suggesting a platform for South-South collaboration. Other causes for the dominance of the international collaborations in scientific communications in WoS are the cognitive dominance of the biomedical and natural sciences, where collaboration among geographical dispersed groups of individuals is very common. The type of research that results in publications indexed in WoS Core Collection might also cause the dominance of international collaboration in WoS when compared to SciELO CI. Researchers from LAC countries might have a marginal participation in these collaboration networks. This position results of a collaboration among many authors and contributions in the form of data processing instead of cognitive contributions and argumentations. Successful collaborations in the region should hold the researchers in leadership positions (Moya Anegón et al., 2013).

From a cognitive perspective, the inclusion of SciELO CI into WoS offers new opportunities of coverage of disciplines and specialties where the particularities of the territory and the social context are important. Public health, social sciences and agriculture are relevant in SciELO CI; the participation of the LAC scientific communications in these disciplines in the core collection of the WoS has traditionally been low. In this sense, the 15% overlap of SciELO CI journals in both indexes suggests that the inclusion of SciELO CI in the WoS benefits WoS in terms of coverage of regional scientific advances, particularly of communications that have a local object of study and where communication is more original and responds to regional capabilities, but also regional issues and problems.

The inclusion of SciELO CI has raised some concerns among the editors of Spanish⁴ and Portuguese journals that have benefitted from a special treatment and inclusion in WoS but that do not have an important position in SciELO CI. Editors of these journals fear that the policy of articulation of SciELO CI into the WoS might result in exclusion of their journals from WoS.

Inclusion of SciELO CI into WoS, responds to the need for a more inclusive representation of scientific results despite regional constraints and conditions. This has resulted from the competition of services offered by Thomson Reuters and Elsevier. The strategies aimed at improving regional visibility are different in Scopus and in the Web of Science. While Scopus has aimed at increasing coverage by increasing their base of regional journals, the globalization of the Web of Science (Testa, 2011) has meant the articulation of regional exercises. The Chinese Journal Database has been hosted in the WoS since 2008, the

⁴ FECyT (Spain's foundation for science and technology) has had an important role in certifying quality of its quality journals in order to support their inclusion in the WoS after an alliance with Thomson Reuters around 2007 (FECyT, 2011)

inclusion of SciELO CI and the Korean Journal Database has been operative since 2014. We believe that the strategy followed by Thomson Reuters provides the cumulative expertise of circulation and visibility promoted regionally, by programs similar to SciELO. We would like to explore this issue further in the future to understand how the inclusion of SciELO CI might put the WoS back in the competition for visibility of regional results.

References

- Aguillo, I. (2014). Políticas de información y publicación científica. *El Profesional de la Información*, 23 (2), 113-118.
- Arunachalam, S. & Doss, M.J. (2000). Mapping international collaboration in science in Asia through co-authorship analysis. *Current Science*, 79 (5), 621-628
- Chandiwana, S. & Ornbjerg, N. (2003). Review of North South and South South cooperation and conditions necessary to sustain research capability in developing countries. *Journal of Health Population and Nutrition*, 21(3), 288-97.
- Elsevier. (2007). Elsevier News America Latina. Retrieved on January 10, 2015 from: http://www.elsevier.com.br/bibliotecadigital/news_dez07/pdf/edicao_03_esp_ok.pdf
- FECYT. (2011). Análisis de la presencia de las revistas científicas españolas en el JCR de 2010. Retrieved on January 10, 2015 from: http://icono.fecyt.es/informespublicaciones/Documents/2011_07_27RevEspanolasJCR2010.pdf
- Garfield, E. (1971). The mystery of the transposed journal lists—wherein Bradford's Law of Scattering is generalized according to Garfield's Law of Concentration. *Current Contents*, 3(33), 5–6.
- Lucio-Arias, D. (2013). Colaboraciones en Colombia, un análisis de las coautorías en el Web of Science 2001-2010. In, J. Lucio (Ed.). *Observando el sistema nacional de ciencia y tecnología, sus actores y sus productos*. Bogotá: OCyT.
- Leydesdorff, L., & Bornmann, L. (in press). The Operationalization of "Fields" as WoS Subject Categories (WCs) in Evaluative Bibliometrics: The cases of "Library and Information Science" and "Science & Technology Studies". *Journal of the Association for Information Science and Technology*. <http://arxiv.org/abs/1407.7849>.
- Meneghini, R., Mugnaini, R. & Packer, A.L. (2006). International versus national oriented Brazilian scientific journals. A scientometric analysis based on SciELO and JCR-ISI databases. *Scientometrics*, 69(3), 529-538.
- Rafols, I., Porter, A. L. & Leydesdorff, L. (2010). Overlay science maps: a new tool for research policy and library management. *Journal of the American Society for Information Science and Technology*, 61(9), 871–1887.
- Vessuri, H., Guédon, J.C., & Cetto, A.M. (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), 647-665.
- Testa, J. (2011). The globalization of the Web of Science. <http://wokinfo.com/media/pdf/globalwos-essay.pdf>.