Open Access Indicators and Scholarly Communications in Latin America

Juan Pablo **Alperin**
Dominique **Babini**
Gustavo **Fischman**
[editors]

United Nations Educational, Scientific and Cultural Organization
FLACSO BRASIL
PKP PUBLIC KNOWLEDGE PROJECT
SciELO
redalyc.org
CLACSO

OPEN ACCESS INDICATORS AND SCHOLARLY COMMUNICATIONS IN LATIN AMERICA

Open Access Indicators and Scholarly Communications in Latin America / Juan Pablo Alperin ... [et.al.] ; edición literaria a cargo de Juan Pablo Alperin ; Dominique Babini ; Gustavo E. Fischman. - 1a ed. - Ciudad Autónoma de Buenos Aires : CLACSO; Buenos Aires: Unesco, 2014.
E-Book.

ISBN 978-987-722-027-8

1. Educación. 2. Comunicación. I. Alperin, Juan Pablo II. Alperin, Juan Pablo, ed. lit. III. Babini, Dominique, ed. lit. IV. Fischman, Gustavo E., ed. lit.
CDD 302.23

Otros descriptores asignados por CLACSO:
Acceso abierto / América Latina / Comunicaciones académicas / Indicadores acceso abierto / Repositorios digitales / Evaluación de la investigación / SciELO / Redalyc / PKP / CLACSO

Other descriptors allocated by CLACSO:
Open Access / Latin America / Scholarly communications / Open access indicators / Digital repositories / research evaluation / SciELO / Redalyc / PKP / CLACSO

OPEN ACCESS INDICATORS AND SCHOLARLY COMMUNICATIONS IN LATIN AMERICA

Juan Pablo Alperin, Dominique Babini and Gustavo Fischman
(Editors)

Juan Pablo Alperin
Abel L. Packer
Eduardo Aguado-López
Arianna Becerril-García
Dominique Babini
Gustavo Archuby
Valeria Carrizo
Darío A. García
Sebastián Higa
Diego Spano

Consejo Latinoamericano de Ciencias Sociales
CLACSO
Conselho Latino-americano de Ciências Sociais
Latin American Council of Social Sciences

Deputy Executive Secretary Pablo Gentili

Academic Director Fernanda Saforcada

Editorial Production & Web Contents Area

Chief Editor Lucas Sablich
Art Director Marcelo Giardino

Production Fluxus Estudio

First edition
Open access indicators and scholarly communications in Latin America (Buenos Aires: CLACSO, august de 2014)

ISBN 978-987-722-027-8

CLACSO
Consejo Latinoamericano de Ciencias Sociales - Conselho Latino-americano de Ciências Sociais (Latin American Council of Social Sciences)
Estados Unidos 1168 | C1101AAX Ciudad de Buenos Aires, Argentina
Tel. [54 11] 4304 9145 | Fax [54 11] 4305 0875 | <clacsoinst@clacso.edu.ar> | <www.clacso.org>

Sponsored by the Swedish International Development Agency Sida

TABLE OF CONTENTS

FOREWORD

THE WORLD TODAY BOASTS phrases such as "death of distance" suggesting that distance is no longer a limiting factor in people's ability to interact and communicate. Another aphorism is that the world is "flattened" in terms of opportunities, which is seen as being facilitated by advancement of Information and Communication Technologies (ICTs) that have allowed convergence of knowledge pools and resources from all over the world.

Today, from the contexts of knowledge acquisition, knowledge processing, knowledge dissemination and knowledge utilization, the fundamental questions of what, where, why, whom and when have all changed. Evidently, the key ingredient that is affording this change lies in the way and pace with which science and technology (S&T) are advancing and "opening up". Indeed, S&T have created enormous new potential for information and knowledge exchange and for the emergence of new patterns of communication and exchange of ideas. Thus science and technology and their inherent "openness" are increasingly recognized as fundamental to creating sustainable knowledge societies. Not least to note "A successful modern economy is founded on a strong [and open] scientific base that has the ability to convert

scientific research and knowledge into products and services, which bring social and economic benefits[1]".

In developing regions, a high percentage of scientific research results are published in local languages. Also most of these publications serve a limited geographical space and scale. Hitherto, only a very small fraction of these publications are included in the so-called "international" indexes that produce the most widely used indicators to evaluate individual and institutional research outputs of the developing regions.

The Internet and the Web have provided developing regions with unique opportunities to build regional initiatives that give visibility and free access (open access) to local and regional research publications. Initiatives in Latin America have shown outstanding results in the past decades and are now matured enough to expand their influence as complementary or alternative sources of indicators for the evaluation of research outputs.

As regional open access initiatives from Latin America have begun to produce and share indicators, this book on *"Open Access Indicators and Scholarly Communications in Latin America"* is a first attempt to systematically explore and describe them within the broader context of Open Science. In this process, the publication aims at identifying the presence, growth, use and reach of research results that are now "open" and freely available in the Web.

This book is the result of a joint research and development project supported by UNESCO and undertaken in 2013 by UNESCO in partnership with the Public Knowledge Project (PKP), the Scientific Electronic Library Online (SciELO), the Network of Scientific Journals of Latin America, the Caribbean, Spain and Portugal (RedALyC), Africa Journals Online (AJOL), the Latin America Social Sciences School-Brazil (FLACSO-Brazil), and the Latin American Council of Social Sciences (CLACSO). This book aims to contribute to the understanding of scholarly production, use and reach through measures that are open and inclusive. This publication serves as an important milestone in UNESCO's Open Access Strategy that 196 countries have collectively endorsed. The publication has also received significant inputs from the PKP-FLACSO-IDRC's project "Quality in Open Scholarly Communication of Latin America", which was undertaken in 2012-2013 in collaboration with Latindex, SciELO and Redalyc.

---

1 Creating a knowledge-based society"; An interview with Noel Treacey, Minister for Science, Technology and Commerce of the Republic of IrelandEMBO Rep. 2000 December 15; 1(6): 460–462.

The present book is divided into two sections. The first section presents a narrative summary of Open Access in Latin America, including a description of the major regional initiatives that are collecting and systematizing data related to Open Access scholarship, and of available data that can be used to understand the (i) growth, (ii) reach, and (iii) impact of Open Access in developing regions. The first section ends with recommendations for future activities. The second section includes in-depth case-studies with the descriptions of indicators and methodologies of peer-review journal portals SciELO and Redalyc, and a case of subject digital repository maintained by CLACSO.

We hope that this book will provide some headway in introducing inclusiveness in the global Open Access scholarly communications and will ensure full participation of an important developing region on the world. We also believe that this book will help appreciate the contribution that Latin America is making in rethinking the way the world is producing, distributing, using and evaluating research output in the digital world.

We would like to take this opportunity to thank the editors: Dr. Juan Pablo Alperin; Dr. Dominique Babini and Prof. Gustavo Fischman and Project Coordinator Dr. Bhanu Neupane for their hard work in bringing the book to a shape that it bears now.

Dr. Pablo Gentili
Executive Secretary
CLACSO-Latin American
Council of Social Sciences

Dr. Indrajit Banerjee
Director
Knowledge Societies Division
UNESCO

SECTION I

OPEN ACCESS INDICATORS

ASSESSING GROWTH AND USE OF OPEN ACCESS RESOURCES FROM DEVELOPING REGIONS

The Case of Latin America* **

Juan Pablo Alperin***

1. INTRODUCTION

Research in developing regions is almost exclusively government funded, primarily through national or federal state agencies and to a

* With the collaboration of Dominique Babini (CLACSO) & Gustavo E. Fischman (ASU/FLACSO-Brazil).

** This chapter was produced as part of the UNESCO-funded project: *Open Access Indicators: Assessing Growth and Use of Open Access Resources from Developing Regions—The Case of Latin America*. Juan Pablo Alperin (Public Knowledge Project/Stanford University) was the lead author of this chapter. Dominique Babini (CLACSO) and Gustavo E. Fischman (Arizona State University/FLACSO-Brazil) collaborated in the conceptualization and development of the initial draft report and the final version of the report. The following organizations each contributed an indicators report and, in some cases, data to inform this work: the Scientific Electronic Library Online (SciELO); the Network of Scientific Journals of Latin America, the Caribbean, Spain and Portugal (RedALyC); and, the Latin American Council of Social Sciences (CLACSO). We would like to thank Bhanu Neupane (UNESCO), Eduardo Aguado Lopez and Arianna Becerril (RedALyC), Abel Packer and Fabio Batalha (SciELO), and Ana María Cetto, Alonso Gamboa and Guillermo Chavez's (Latindex) for their contributions to this chapter. Together, the contributions of these organizations and individuals made this chapter possible. Details on the project can be found online at: http://flacso.org.br/oa/category/proyectos/.

*** Publishing Studies and Public Knowledge Project, Simon Fraser University, Canada.

lesser – but still relevant – extent by international cooperation agencies (UNESCO, 2010). Another salient characteristic is that an important part of the research output is published in local and regional journals. These two patterns – dominant national funding and localized publications – are contrasted with the limitation that the most widely used and accepted scholarly indicators to evaluate research output and impact in developing regions are provided by so-called "international" indexes (mainly Thomson-Reuters' Web of Science [WoS] and Elsevier's Scopus), where only a small fraction of journals from developing regions are included. The shortage of indicators that cover journals from developing regions severely hinders the ability to evaluate their individual and institutional research output in terms of quality, impact, and reach.

With varying degrees across disciplines and countries, a high percentage of research from developing regions is published in local languages, in local and regional journals, research reports, books, and theses. While a selected group of researchers mainly working in the experimental, natural, and biomedical areas have more presence in international journals, other areas where researchers are part of a national and regional conversation about local and regional problems are naturally more localized, as is the case for agriculture, health, and social sciences. Take, for example, the case of the social sciences in Argentina, where the percentage of articles published in local and regional journals is as high as 80% (Gantman, 2011). In Latin America more broadly, over 50% of the journals identified by Latindex (a Directory of journals published in Latin America) are from the social sciences.

Despite strong criticisms against the use of Thomson-Reuters' Journal Impact Factor (IF) to evaluate research output in general (see the San Francisco Declaration on Research Assessment [DORA], 2012),[1] and criticism specific to its use to evaluate research productivity outside of Europe and North America (Alperin et al., 2011; Gúedon, 2008; Vessuri et al., 2013), the IF continues to be widely used to evaluate individual and institutional research outputs in developing regions due to the lack of a viable alternative. Other indicators, like those of SCImago, are being developed based on a larger set of journals from Scopus that, while more inclusive than Thomson-Reuters', continue to cover only a fraction of the research outputs produced by developing regions (Alperin, 2014).

As a result, there is a pressing need to find alternative indicators of research production that capture a larger portion of the outputs from

1 http://www.ascb.org/dora/

developing regions. This chapter investigates the potential of alternative sources of indicators, with a focus on those stemming from Open Access (OA) initiatives. The focus on OA stems from the widespread use of this model of publication in developing regions—especially Latin America— and from the desire to contribute to our understanding of scholarly production, impact, and reach through measures that are open and inclusive of developing regions. Unfortunately, while it is evident that the degree of adoption of OA models is fairly extensive, there are no exact figures. This chapter seeks to delve further into what can currently be inferred from available sources, as well as make recommendations on what could be improved going forward.

Estimates on the extent of OA in developing regions vary significantly. In Latin America, these estimates range from a low of 51% of all online journals being OA to a high of 95% depending on the source of the data used. In the Scopus database, 74% of all Latin American journals are OA, compared with their global total of 9% (Miguel et al., 2011). In the Ulrich's Periodicals database, 51% of all online journals were found to be OA, compared with their global total of 7% (Haider, 2005). The highest estimate, although not based on a rigorous study, comes from the director of the Scientific Electronic Library Online (SciELO), an expert in scholarly communications in Latin America, who suggests that 95% of all online journals published within the region are fully OA (Abel Packer, personal communication).

Unfortunately, none of the databases that collect subscription information provide an adequate sample from which to gather a more exact estimate. But, even these varied estimates suggest much higher levels of OA participation than other regions of the world, especially in Latin America, at least in comparison to the global totals in the studies cited above and to other global estimates of OA, which are estimated around 20% (Laakso & Björk, 2012; Laakso et al., 2011).

More globally, 48% of the DOAJ's 9,137 journals are from developing countries, and an estimated 4,000 of the approximately 6,000 journals using the Public Knowledge Project's (PKP) Open Journal Systems (OJS) are from developing regions (PKP, 2013)[2]. Furthermore, an estimated 83% of these are fully OA (Edgar & Willinsky, 2010). These estimates provide a clear indication that this is a distinctive feature of scholarly communication in developing regions.

Given this extensive adoption of OA, and more than a decade of successful OA initiatives operating in developing regions (Babini, 2012, 2013; GOAP, 2013; UNESCO, 2013), it is now feasible and urgent to explore their data and build indicators that complement the

2 http://pkp.sfu.ca/ojs/pkp-stats

traditional indicators, which so poorly represent developing regions, with ones that are better suited to the realities and needs of developing regions. In this chapter, we explore these data to identify indicators of the presence, growth, impact, and reach of OA. In this sense, the term "OA indicators" has two meanings: indicators *about* OA and indicators *derived from* OA initiatives.

This chapter seeks to guide the reader to an understanding of the existing and desired OA indicators through the cooperation of major OA initiatives from Latin America. In past years, a few studies have approached the issue of OA indicators in general and in Latin America (Alperin, 2011; Gómez et al., 2009; Mayr, 2006; Villanueva, 2013). While we draw on these studies, this chapter seeks to more systematically document existing and desired indicators.

The report is divided in four main parts. After a brief description of the scope and definitions, the first major section provides an overview of the available data on OA from selected developing regions, including a description of the major regional initiatives that are collecting and systematizing data related to OA scholarship in Latin America, newer alternative indicators, and the well-established bibliographic databases. In the three sections that follow, the available data is described along with the current state of indicators that can be used to understand the (i) *growth*, (ii) *reach*, and (iii) *impact* of OA in developing regions. Naturally, these three sections draw on the available data described in the first part, but delve into the specific indicators that can be used for each purpose. Each section ends with some recommendations on the type of activities that we believe should be adopted, supported, or continued.

2. SCOPE AND DEFINITIONS

The OA community has recently affirmed the definition of OA introduced in the original Budapest Open Access Initiative from 1993:

> By "open access" to [peer-reviewed research literature], we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited.[3]

3 http://www.webcitation.org/6AdLcI1TC

While this is likely the most accepted definition, others vary on the specific materials to which the definition applies. For example, the Berlin Declaration speaks of OA contributions applying to "original scientific research results, raw data and metadata, source materials, digital representations of pictorial and graphical materials and scholarly multimedia material".[4] In contrast, the Salvador Declaration on Open Access, perhaps the most relevant in the context of this chapter as it focuses specifically on the developing world, speaks of access and use of "scientific information" (defined as scholarly works).[5] It is around these definitions that the scope of this chapter is based, seeking to explain and understand OA indicators specific to content published in peer-reviewed journals and, to a lesser degree, institutional repositories.

For the authors of this chapter, however, it was important to explicitly acknowledge that even though the formal definitions of OA used in these declarations are implicitly or explicitly focused on peer-reviewed literature and related materials, OA could be applied more broadly. In developing regions, such as Latin America, a very significant portion of research-related activities is in fact disseminated outside the system of peer-reviewed publications. Research centers and non-governmental organizations, for example, produce a great deal of literature (i.e., reports, program evaluations) that is arguably more relevant for the needs of the communities in developing regions. Tragically, there is little systematization of this information and even fewer indicators to track its growth, use, or impact, even in those fields with a tradition of regional subject digital repositories (i.e., agriculture with SIDALC, health with BVS, and the social sciences with CLACSO). It is our strong belief that those contributions should not be ignored, even as they fall beyond the scope of this chapter. We therefore focus on OA indicators of scholarly articles published in academic journals and in subject and institutional repositories.

Within these parameters, this chapter focuses on the indicators of *growth, reach, and impact* of OA. We apply definitions first specified in the research agenda produced by the Quality in the Scholarly Communication of Latin America project.[6]

---

4 http://openaccess.mpg.de/286432/Berlin-Declaration

5 http://www.icml9.org/meetings/openaccess/public/documents/declaration.htm

6 The Latin American institutions collaborating in the production of this chapter set out to establish working definitions to be used when conducting research, but they are equally useful for framing and understanding the use of indicators related to OA. These definitions are not presented here in a normative way – each surely varies depending on the context – but they are the definitions

By *growth*, we refer to increases in the number of scientific publications and scholarly publication venues. This differs from the growth of the size of various databases and services, which is all too often mistaken for growth in the underlying phenomenon. Tracking the actual growth of scientific publications and scholarly publication venues is of particular relevance, especially as research and research communication gain more importance in the 21st century educational landscape (Alperin, 2011).

By *reach* we mean the extent and diversity of channels used to access scientific journals and scholarly articles, their various venues, and the public that accesses them. In this sense, reach refers to various demographic characteristics of the readership of scholarly content, including descriptions of the channels where the research is disseminated and information on the people who access the content. This definition attempts to disentangle reach from impact.

In contrast to reach, by *impact* we mean the degree to which research published in scholarly journals is *applied*, *cited*, *discussed*, or has otherwise *somehow affected* individuals and or groups within or beyond academic communities. That is, we draw a distinction between access (i.e., reach) and use (i.e., impact), with the former being a prerequisite for the latter. Measuring impact is surely elusive (especially social impact), but this chapter applies this broad definition and calls attention to the forms of impact that can most readily be measured.

3. AVAILABLE DATA

Although Latin America has been characterized in this chapter under the label of "developing regions," there are large differences between it and other developing regions in the realm of scholarly communications. While developing regions in general have made great strides in their scholarly communication strategies in the last decade, Latin America is unique in the world in its adoption of OA models in communication. This adoption has led to the development of several parallel, regional initiatives all working to increase access to research published in Latin America and simultaneously improve the quality of the region's scientific journals. As a result of these initiatives, the development of OA indicators in the region is also years ahead of other parts of the world.

---

used in the production of this chapter and thus serve as a lens through which to understand the choice of indicators and recommendations made herein. For details on the project, see http://flacso.org.br/oa/calidad-en-la-comunicacion-cientifica-abierta- de-america-latina/. For the research agenda, see http://dx.doi.org/10.6084/m9.figshare.847295

While regions rely on the Directory of Open Access Journals (DOAJ)[7] for tracking OA journals, Latin America has more comprehensive databases (more detail on this below). Authors such as Björk, Laakso, and their colleagues (e.g., Björk, 2011; Björk et al., 2010; Laakso & Björk, 2012; Laakso et al., 2011) have used survey methods in an attempt to track the number of OA articles published globally over time. They estimate that just under 20% of articles are published in OA journals globally, a valuable point of comparison for understanding the development of OA in developing regions. However, while their work is extremely innovative, it is also very labor intensive and proves inadequate for ongoing or targeted purposes.

Their estimate is complemented by a study at the journal level using data from Scopus (Miguel et al., 2011). While Scopus is the largest single-source article-level index in the world, its content, like that of the other major commercial index (Thomson-Reuter's WoS) does not, by design, accurately reflect the output of developing regions. These two commercial/scientific databases are highly skewed towards publications from traditional scientific commercial publishing in the United States and Western Europe, to the detriment of publications from other regions, languages other than English, and emerging OA models. Thomson-Reuters even explicitly acknowledges in its selection criteria that the Web of Science "focuses on journals that publish full text in English, or at very least, bibliographic information in English" (Testa, 2012, n.p.), citing that "English is the universal language of science" (Testa, 2012, n.p.) It is beyond the scope of this chapter to discuss to what extent English is currently the scientific language, but there is no doubt that in Latin America, as well as in many scientific disciplines, there are many publications in national languages.[8]

The privileging of English as the language of publication is one of the many biases of the commercial scientific databases focused on the global North, and undoubtedly Scopus and Web of Science have been the key players in establishing standards of quality, and their data and resulting indicators have been well established. While these indicators have the benefit of decades of experience of prominent commercial global companies, they have been unevenly adopted by many scientists, universities, and research centers in the global South. As we will see below, both of these commercial databases severely underestimate the scholarly production of the region and provide a skewed and misleading picture of the publishing activities of developing countries.

---

7 http://www.doaj.org

8 According to Latindex, there are 13,446 Spanish language journals and 5,297 Portuguese language journals in Latin America.

Voices clamoring for creating alternatives to Scopus and WoS are increasing,[9] but generating reliable and compelling data will take a great deal of effort. As will become evident in this chapter, methodologies and established best practices around OA indicators are still lacking, but there are also valuable experiences in the production of OA indicators from developing regions, in particular Latin America.

The existing indicators primarily stem from regional initiatives that generate locally curated data sets, applying local standards to determine what is included (unlike WoS and Scopus, which impose criteria on developing regions). We therefore begin our exploration of the available data from these traditional databases. This is followed by richer data from the OA initiatives in developing regions. We subsequently present other sources of data, including what are known as "altmetrics," and describe the process by which journals and research are included in the presented initiatives.

3.1. TRADITIONAL SOURCES

3.1.1. Web of Science and Scopus

Perhaps the most well known source for bibliographic and bibliometric data is Thomson-Reuters' WoS. This database has been a key reference globally and used for such analysis for several decades. This is not surprising, given that Eugene Garfield, the creator of the Science Citation Index in 1963 (a predecessor to the WoS) repeatedly suggested in his early work and elsewhere the citation index is sufficient to understand scholarly production from around the world (Garfield, 1983a, 1983b, 1996; Garfield & Sher, 1963). However, we believe as Cetto and Alonso-Gamboa (1998) do, that "it is difficult to justify…the use of [WoS] figures as official statistical indicators of national scientific productivity, or as indicators of performance" (p. 92).

As Cetto and Alonso-Gamboa (1998) point out for case of Latin American, databases like the WoS are biased against publications from developing regions. The bias can be seen in Figure 3.1, which shows in a dramatic way the differences in the representation of various regions in the world by scaling a country's size in proportion to the number of journals included in the WoS in 2010, where there are only 242 Latin American journals – more than double from 2006 after the WoS sought to expand its international coverage (Testa, 2011). The argument for this bias has always been that of "mainstream" or

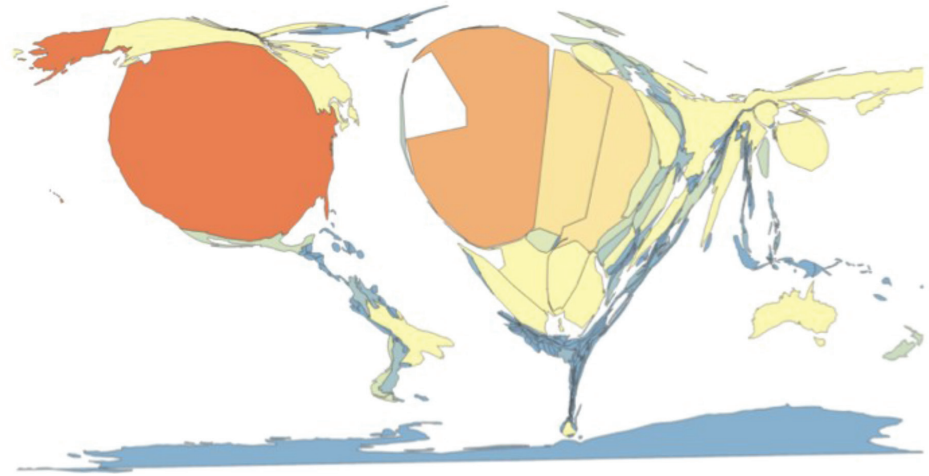
9 DORA (2012) is probably the best example, with over 10,000 individual signatures and nearly 500 institutions declaring the need for alternative forms of research assessment.

"international" science, and although many take exception to this argument, it is irrelevant here. The end result, regardless of its rationale, is that the WoS is an inadequate dataset to study scholarly communications from developing regions.

Aside from the evident issues of equity, the underrepresentation of developing regions in these datasets severely limits conclusions on scholarly productivity and impact. Extracting meaning of any analysis done at the disciplinary level within that limited subset is surely flawed. After all, the 242 journals listed in WoS' 2010 edition represent only 4% of all journals in the WoS and only 3% of the known entries in the Latindex Catalog.

Thus, the coverage by field within the already limited number of journals from developing regions is quite sparse. As Moed (2009) points out, in "fields with a moderate coverage, language or national barriers play a much greater role than they do in other domains of science and scholarship" (p. 14). Thus, analyses are subject to distortions caused by all the missing entries in the dataset. The WoS may be an effective tool for identifying trends and patterns of the core set of journals represented, but for journals who are not part of that core set it is impossible to say if any trends or patterns reflect the vast majority of journals from developing regions that are not included in the analysis.

Figure 3.1.
Cartogram of Number of Journals in WoS in 2012



The coverage in Scopus is decidedly better, with more than double the journals from Latin America than the WoS, but its relative coverage is still highly biased towards North America and Europe. Scopus currently indexes nearly 600 journals from Central and South America. In science-related fields, Scopus indexes upwards of 97% of articles and reviews indexed in the WoS (that is, the WoS is a subset of Scopus in these fields) (Moed, 2009). For other fields, however, the coverage varies and any resulting indicators would differ between the two sets, especially of fields that do not have adequate representation (Bakkalbasi et al., 2006; Meho & Yang, 2007).

Given the differences in coverage and the commonality of underrepresentation of developing regions, the need for a "comprehensive and reliable information system that duly gives account of the production of scientific periodicals [from Latin America]" (Cetto & Alonso-Gamboa, 1998, p. 91) becomes evident.

3.1.2. Ulrich's Web Global Serials Directory

Ulrich's Web Global Serials Directory tout's itself as "the world's most authoritative source of bibliographic and publisher information for serials and the most comprehensive source of print and electronic serials data available".[9] Like the WoS and Scopus, the data from Ulrich's is only available by purchasing a subscription. The data includes basic bibliographic details, including International Standard Serial Number (ISSN), volume, issue number, and country of publication. It also very usefully tracks whether or not a publication is OA.

Of the commercial indexes, Ulrich does have the more comprehensive coverage of Latin American publications, with 942 Latin American journals resulting from a search of active, academic/scholarly, peer-reviewed journals in the online version of the Knowledgebase (as of September 2013).[10] Again, this represents a mere 1% of the total journals covered, indicating the inherent bias against journals from developing regions and the inadequacy of this set for understanding OA in developing regions.

3.1.3. Directory of Open Access Journals

Unlike the previous databases mentioned, the DOAJ exhibits a strong presence of journals from developing regions. This stems from the DOAJ's mission to be "comprehensive and cover all open access scientific and scholarly journals that use a quality control system to guarantee the content."[10] This selection criteria is not based on a fixed notion of a journal's "impact," but rather on a basic set of criteria around the

10 http://ulrichsweb.serialssolutions.com/

type of content (focused on research), target audience (researchers), and type of access (open, with no embargo period). DOAJ does select on quality, but quality is defined only by a journal's self-reporting of an "editor, editorial board and/or a peer-review system."[10] These criteria are inclusive of research from all regions of the world.[11]

There are currently 1823 Latin American journals out of 9,746 (19%). These greater numbers are partially a reflection of the more equitable inclusion practices, and partially a result of the greater levels of OA found in developing regions. DOAJ provides description at the journal-level and, for some journals, at the article-level when the data can be harvested using established standards and protocols (as of November 2013, article-level metadata was available for 5,673 of the 9,746 journals).[12]

The DOAJ serves as an easy-to-use and readily available dataset on OA journals from developing regions. Thanks to cooperation between DOAJ and the OA initiatives, DOAJ has complete information on OA journals and thus its content is analyzed directly through the initiatives described in more detail below.

3.2. OPEN ACCESS INITIATIVES

In Latin America, three major non-commercial, publicly-sponsored OA initiatives – Latindex, SciELO, and RedALyC – provide the best available data on the scholarly publishing activities of the region (arguably of any region). The three combined datasets provide a reasonable approximation of the state of scholarly journals in Latin America in a way that no single database could, especially not those traditionally used for bibliometric analysis, such as WoS or Scopus. Together, they have jointly defined standards for what constitutions a scholarly journal and one of them, Latindex, has used a cooperative approach to collecting bibliographic information on serial publications that meet those standards.

3.2.1. Latindex

Latindex is the most inclusive and comprehensive source of information regarding scholarly journals in Latin America, the Caribbean, Spain, and Portugal. Established in 1997, it is also the oldest of the three major scholarly communication information systems working at the regional level in Latin America. Through cooperation with in-

11 More detail on the selection criteria can be found on DOAJ website: http://doaj.org/about#criteria.

12 Details on which journals are available at the article-level are not directly available through the data file provided by DOAJ.

stitutions in 22 member countries (including all Spanish-speaking countries in Latin America), Latindex maintains both a Directory and a Catalog of periodical publications.

The Latindex Directory contains bibliographic information of over 20,000 periodical publications (approximately 5,000 are from Spain and Portugal). Of the 15,261 from Latin America and the Caribbean, 13,041 are known to be actively publishing, with another 677 of unknown status and 1,543 that have ceased publishing. In Latin America alone, there are currently 5,408 active journals in the Latindex Catalog (out of 5,665 entries). These journals can be broadly categorized as scientific or research journals, professional and trade publications, and science and cultural dissemination magazines (i.e., popular science magazines). The Directory aims to be all encompassing by not applying any editorial criteria for inclusion, other than requiring the contents be of "academic interest." In contrast, the Latindex Catalog is a curated list of periodicals that meet a minimum set of criteria that help determine the quality of the publication.[13] While there are gaps in the coverage – and it is impossible to verify if all countries update their corresponding records with equal accuracy, completeness, or speed – the Latindex Catalog is the closest available proxy for the universe of Latin American publications meeting a basic set of editorial criteria.

The Latindex model has proven effective for ensuring representation of journals from all countries in the region, even countries with small or weak scientific publishing systems. Membership and representation to Latindex is decided by a team of scholars, researchers, and scientific organizations from each country, typically a representative from the national science council (or equivalent body). This decentralized approach brings the data collection and evaluation as close to the local source as possible. Well-defined standards and annual meetings ensure that the evaluation criteria are applied equally across publications and countries and concerns are addressed.

Latindex aims to provide *comprehensive* coverage of the regional scientific output, and the inclusiveness of the Latindex Directory is its greatest strength. While it cannot ensure complete coverage, it can approximate it by delegating the responsibility of finding publications and evaluating them to its members, who are situated in each of the participating countries.

For each entry in the Directory or Catalog, Latindex collects a long list of metadata fields (Table 3.1). Not all of these fields are publicly

13 For a full list of the selection criteria, see http://www.latindex.unam.mx/latindex/catalogo.html

visible, although this may change with the launch of a new version of the website in late 2014. Some of these are also available through a public API. Unfortunately, not all fields are complete for all publications and some preliminary analyses have shown there to be some inconsistencies with the data. Nevertheless, the available metadata provide a useful starting point for analyzing scientific and scholarly publications in the region.

3.2.2. SciELO

SciELO operates as a set of national collections of scholarly (i.e., peer-reviewed) journals using a common set of editorial criteria and methodologies, and a common software platform. SciELO has been in operation since 1998, but provides content dating back to the early 1900s (although most of the content is from the last 10 years). It currently operates 16 national collections and three thematic collections, with over 1,100 journals and over 450,000 research articles and other document types. Each collection is operated independently and has its own portal, although a consolidated search function and centralized web services are available on their website.[14]

SciELO publishes and indexes full-text articles for all of their journals and has gained a special place in the region and beyond because they offer citation data for every published article in the platform.[15] SciELO methodology requires that articles be semantically coded in Extensible Markup Language (XML). As a result, SciELO provides the most complete citation data for Latin American publications.

A regional database is of particular importance because most of Latin America's publications are still published in local languages and often cater to, and cited by, a more local audience. This is not the case for all regional publications, but it is certainly the case for some (Meneghini et al., 2006). It is therefore of particular importance to capture citations from other Latin American publications, something that cannot be done in other citation databases.

---

14 http://www.scielo.org

15 SciELO and RedALyC both publish and index full-text articles for all of their journals and, unlike Latindex, provide information at the article level, not just the journal level. While SciELO, as well as RedALyC, are technically publishers (that is, they "publish" content online that in some cases has not been made available elsewhere), they are not publishers in the traditional sense (i.e., they are not involved in the operation of the journals, do not provide financing, editing services, or any other oversight of the editorial or production process). They have also been described as a "meta-publishers" (Packer & Meneghini, 2007) and as a hybrid between a repository and a publisher (Gúedon, 2008).

Table 3.1
Fields available in Latindex Catalog

Information	Available online
Unique journal ID	
Journal format ID	
Date of inclusion	
Late modified date	
Type of register	
Title	✓
Abbreviated title	✓
Previous titles	
Subsequent titles	
Country	✓
Status	✓
Start date	✓
End date	✓
Frequency of publication	✓
Type of publication	
Format	✓
Language	✓
ISSN	✓
ISSN2	
Subjects	✓
Dewey universal classification	
Dewey classification	
Responsible entity	
Place of publication	✓
Publisher	✓
Contacts	
Address line 2	
Address line 2	
City	✓
Province/state	✓

Information	Available online
Country of editor	✓
ZIP code	
E-mail	
URL	
Phone number	
Fax number	
Indexed in	✓
Type of publication	✓
Type of responsible entity	✓
Price	
Distribution (run)	
Distribution (form)	
Distribution (channels)	
Distribution (geographically, national)	
Distribution (geographically, international)	
Distribution (geographically, # of countries)	
Notes	
Full-text URL	✓
Span (time)	✓

The ideal for capturing citations is, of course, to capture all citations, whether they are in local, regional, or international journals.

In October 2013, in partnership with Thomson-Reuters, SciELO launched the SciELO Citation Index[16] to combine its citation data with that of journals from the WoS. As this is a relatively recent initiative, the specifics of the Citation Index, or further details on coverage, are not yet available. The first annual report with citations from both SciELO and WoS journals is expected to be available in July 2014, and it should include a more expansive set of indicators than is currently available SciELO alone.

Citation data is difficult and time consuming to extract and each SciELO portal assumes the responsibility and costs of doing this markup. In some cases, it is delegated to the journals them-

16 http://thomsonreuters.com/press-releases/102013/SciELO-Collaboration

selves and in others it is the organization in charge of SciELO that assumes the costs and efforts. The SciELO citation database has grown to over 10 million records as of May 2014. The data are not entirely free of error, but there are processes in place to attempt to automatically correct errors. Further checks and a clean-up of the database are planned in the future, but dependent on funding. It is one of SciELO's main objectives to make available an open database that is as complete and clean as possible for bibliometric and scientometric studies.[17]

SciELO also provides co-authorship indicators in the form of counts of articles authored between any pair of countries. These indicators cannot currently be downloaded wholesale, but they can be located through the online interface by selecting several journals, the country pair of interest, and the years that should be aggregated. It should be noted that SciELO country affiliations are very often missing and therefore this information is often incomplete, especially for the non-corresponding authors.

To complement their bibliometric indicators, SciELO also provides access to download statistics for the journals in the SciELO Brazil and Chile sites. Unfortunately, other collections do not currently have the capacity to host the access logs for all of their content, although there is a project underway to provide access statistics for the entire network. For the journals that have access statistics available, it is possible to view monthly access requests to the journal homepage, article pages (by language or combined), and the table of contents (i.e., issue pages). These statistics are available at the journal level from the sidebar, but also at site and collection levels from the "SciELO in numbers" box on the site homepage.[18]

More details on the indicators offered by SciELO can be found in the chapter produced by SciELO as background for this chapter. It includes details on how to access the indicators and SciELO's future plans for overcoming some of the inherent limitations in the current processes.

3.2.3. RedALyC

A slightly newer indexing and publishing platform is RedALyC, which now publishes over 800 journals from Ibero-America and the Caribbean – hosting over 300,000 articles and other document types. Re-

17 As of early 2014, SciELO released – but had not yet formally announced – the creation of an Application Programming Interface (API) for querying SciELO citations. Details can be found at: http://docs.scielo.org/projects/citedby/en/latest/

18 http://www.scielo.org

dALyC's data is by the far the most complete co-authorship dataset available for Latin America. Like the Latindex Catalog and SciELO, RedALyC selects journals based on a set of criteria that ensures a minimum standard of editorial quality. While RedALyC has also been working on processing citations, the work is not yet complete and at the time of writing there was no projected date for when citation indicators would be ready for the public.

Unlike the other initiatives described above, RedALyC has instead focused on normalizing the metadata related to author affiliations, allowing for accurate information regarding national and international collaborations. The affiliation metadata is available for articles published from 2005–2011 (with 2012 forthcoming) and so-called "production profiles" are available for every journal, institution, and country from the RedALyC website and research lab website.[19] These production profiles contain aggregate percentages of international versus national collaborations, and further breakdown national collaborations between intra- and inter-institutional collaborations. Percentages can be aggregated by institution, country, or discipline, as is shown in several of their already published reports (see De Volder et al., 2013; Gasca Pliego et al., 2013; Rosenzweig et al., 2013). While the raw data is not available for download, the aggregated percentages and counts are available for browsing and sorting through the RedALyC research lab site.

3.2.4. CLACSO

The Latin American Council of Social Sciences' (CLACSO, for its initials in Spanish) Network of Virtual Libraries is a very different type of initiative than those previously mentioned. The goals of the Council, which is a network of 344 research institutions in 21 countries across Latin America and the Caribbean, are the promotion and development of research and teaching in the social sciences. Their network of virtual libraries was set up in 1998 as place for community experimentation and cooperation in the development of OA scholarly communications for members of the CLACSO network. It simultaneously serves to provide international visibility to the social science research community from Latin American and Caribbean, whose production attains scarce visibility otherwise.

Currently, CLACSO's digital repository provides free access to over 30,000 texts, including books, working documents, theses, conference papers, and a peer-reviewed journal collection (hosted on RedALyC), as well as 400 multimedia objects (audio and video) distrib-

19 www.redalyc.org; www.redalycfractal.org

uted among 200 collections (national and thematic) and categorized with over 8000 subject labels.

This service is a decentralized, collaborative work undertaken together with a community of publishers and libraries with a total of 932 participants. Unfortunately, CLACSO does not have dedicated data curation teams like SciELO and RedALyC, and the quality of its metadata has not been subjected to any rigorous evaluations. Still, CLACSO does markup titles, authors, and subject areas (all within the social sciences) and indexes the full-text search of the entire collection.

It should be noted that CLACSO differs greatly from SciELO and RedALyC in that it is not a publisher; it is purely a repository. In that sense, it represents the only "green road"[20] (self-archiving) initiative analyzed in this chapter. CLACSO provides us with just one example, from the social sciences, of the kind of indicators being developed and displayed by a Latin American regional repository that includes peer-review contents as well as research reports, academic books, conference papers, and multimedia, which are the usual contents in digital repositories of the region.[21, 22]

Unlike with journal publishing, where the Latindex Catalog serves as an approximation of all journals published in the region, there is currently no way to estimate what percentage of CLACSO members' output can be found in the repository. Although it is not possible to know how representative the virtual library is of all social sciences production, it does provide a broad window and gives visibility to a wide variety of content produced in Latin America. CLACSO provides statistics on the number of digital objects in the collection as well as access to download statistics aggregated at the collection level, including each of the national collections. As of November 2013, it is also possible to view download statistics at the document level for individual documents and for search results.

---

20 http://en.wikipedia.org/wiki/Self-archiving

21 Other initiatives such as the Alliance of Agricultural Information and Documentation Services of the Americas (SIDALC, for its initials in Spanish) and the Virtual Health Library (BVS, for its initials in Spanish) are as important in their respective realms (agriculture and health) and would have been equally illustrative examples of self-archiving virtual libraries as CLACSO.

22 http://www.opendoar.org/onechart.php?cID=South%20America&ctID=&rtID=&clID=&lID=&potID=&rSoftWareName=&search=&groupby=ct.ctDefinition&orderby=Tally%20DESC&charttype=bar&width=600&caption=Content%20Types%20in%20OpenDOAR%20Repositories%20-%20South%20America

3.3. ALTMETRICS

In recent years there has been increasing interest from the academic and scientific community to develop better indicators that complement the traditional ones typically used for evaluating impact (i.e., citation-based metrics). While there was already a tradition to use a few alternative indicators (most recently usage statistics), as "scholarly communication moves increasingly online, more indicators have become available: how many times an article has been bookmarked, blogged about, cited in Wikipedia and so on" (Piwowar, 2013). These new sources present an opportunity to use metrics that are different (i.e., alternative) to citation-based metrics: hence the name, *altmetrics*.[23] There is no fixed list of sources for these metrics, but they typically include mentions from social media outlets like Twitter and Facebook; links from blogs from sites such as Research Blogging, Science Seeker, and Wordpress.com; citations in Wikipedia; social bookmarking like delicious.com or academic bookmarking like Mendeley, CiteULike, Bibsonomy, and many others. Altmetrics can also refer to metrics on alternative research products, such as presentations, videos, data sets, and software.

In the last few years the field of altmetrics has received a lot of attention as researchers and others explore its potential. For example, Charles W. Bailey (2013) has compiled an altmetrics bibliography; PLOS has a special altmetrics collection[24]; the American Society for Information Science and Technology (ASIS&T) published a special altmetrics issue[25]; and there is a Mendeley group[26] focused on altmetrics. Yet, altmetrics is still in its infancy, especially when compared to the decades old bibliometrics field that it seeks to complement. As such, the meaning, implications, and benefits of altmetrics are still being explored. Although the advantages of altmetrics has not yet been proven, Piwowar (2013, p. 9) points to four potential

---

23 For the sake of clarity, it is important to distinguish between altmetrics and a slightly older but also increasingly common term, "article level metrics" (ALMs). In fact, altmetrics was proposed (in a tweet) by Priem (2010) because he felt that ALM failed to capture the diversity of sources. In contrast, ALMs, which were pioneered by PLOS, are an "attempt to measure impact at the article level [and in doing so] draw from a variety of different data sources, some traditional (e.g., times cited) and some new (e.g., tweets)" (SPARC, 2013). The terms ALM and altmetrics are often used interchangeably, but the distinction is important.

24 http://www.ploscollections.org/article/browse/issue/info%3Adoi%2F10.1371%2Fissue.pcol.v02.i19

25 http://www.asis.org/Bulletin/Apr–13/AprMay13_Piwowar.html

26 http://www.mendeley.com/groups/586171/altmetrics/

advantages that are especially relevant in the context of OA indicators for developing regions.[27]

- A more nuanced understanding of impact, showing us which scholarly products are read, discussed, saved, and recommended as well as cited.
- Often more timely data, showing evidence of impact in days instead of years.
- A window into the impact of web-native scholarly products like datasets, software, blog posts, videos, and more.
- Indications of impacts on diverse audiences including scholars, practitioners, clinicians, educators, and the general public.

As noted above, altmetrics data comes from many sources and takes many different forms. There are currently three major aggregators of altmetrics data that help pull them all together: Altmetric.com,[28] Plum Analytics,[29] and ImpactStory.[30] These services offer a significant overlap in their metrics, although they differ in their approaches to collecting the data and the communities they are trying to serve.

As a result, data from existing sources, as well as virtually all research studies that currently exist using altmetrics data, are being carried out with data from the global north.[31] This situation has prompted researchers to ask whether or not altmterics are propagating global inequality (see Alperin, 2013a). Fortunately, steps are being taken to offer altmetrics services (and as a result collect altmetrics data) in developing regions. For example, Altmetric.com has recently signed an agreement with SciELO Brazil,[32] while PKP has recently launched its own ALM/altmetrics service for journals publishing using PKP's

---

27 Note that Piwowar's use of the term "impact," in this context, actually encompasses both reach and impact under the definitions presented in this chapter.

28 http://support.altmetric.com/knowledgebase/articles/83335-which-data-sources-does-altmetric-track

29 http://www.plumanalytics.com/metrics.html

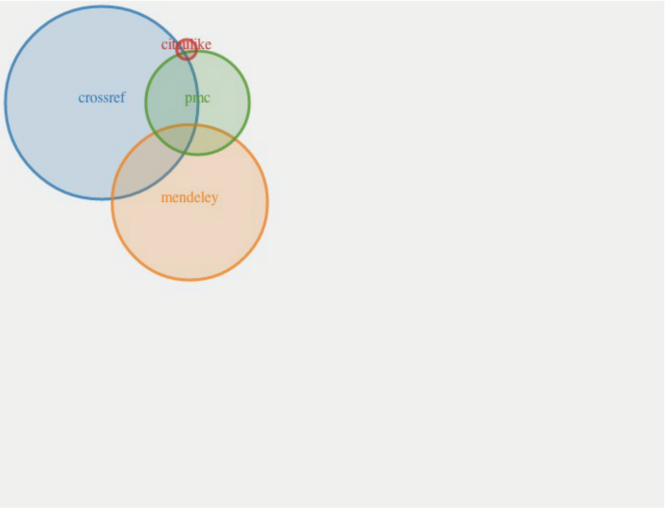
30 http://impactstory.org/faq

31 The OA publisher PLOS has also developed an open source application that it uses for its own journals, which can also be used for a "do-it-yourself" approach by publishers or others. Unfortunately for developing regions, all of the services are reliant on the presence of Digital Object Identifiers (DOIs), which are not very prevalent in developing countries. DOIs are not strictly necessary for the PLOS tool, but the tool relies on another identifier in the absence of a DOI.

32 http://blog.scielo.org/en/2013/08/29/interview-with-euan-adie-ceo-of-altmetric-com/#.UrJ3JGRDtEw

OJS using PLOS' tool.[33] However, neither initiative has yet yielded sufficient data for analysis.

Figure 3.2
ALM Coverage of SciELO Brazil Articles Published in 2012



crossref	1522 (8.5%)
mendeley	989 (5.5%)
pmc	439 (2%)
citeulike	16 (0%)

A very preliminary look at the presence of SciELO Brazil articles from 2012 in a few sources, however, has shown that the level of signal (presence of articles in the new sources) is low. Of over 17,000 articles published in SciELO Brazil in 2012, only 13% had at least one of the collected metrics[34] (Alperin, 2013c). More details on the coverage found can be seen in Figure 3.2 (from Alperin, 2013c).

Although currently there are few altmetrics available for journals from developing regions, services like PKP's and SciELO's relationship with Altmetric.com will begin to grow the available data.

3.4. OTHER SOURCES

There are a number of other indexing services that contain bibliographic information about publications from developing regions – some global in scope (with likely underrepresentation of develop-

33 http://pkp.sfu.ca/pkp-launches-article-level-metrics-for-ojs-journals/

34 See list of sources here: http://pkp-alm2.lib.sfu.ca/sources

ing regions) and others specific for the region. These sources are generally subject-specific, and none provide data other than the bibliographic details and a text-based search of metadata and abstracts.[35] While these are useful services for content discovery specific to the region, it is uncertain what coverage they have beyond the previously mentioned databases.

Working on a global scale, it is important to acknowledge the contribution of Google Scholar in the indexing of information from developing regions. Google Scholar was the first major scholarly index to indiscriminately include works from developing regions, placing them side-by-side with work from the rest of the world. Moreover, Google Scholar provides information regarding citations to journals that would not have it otherwise through "cited by" counts on search results and more recently in "Scholar Metrics." While researchers and journals often use these statistics and counts to evaluate their citation impact, the data is not openly available. Nor is it possible to arbitrarily query the information contained in the Google Scholar index. For this reason, and for this reason alone, it is not included in our analysis.

There are also a growing number of journal portals cropping up around the world. In essence, universities are opting to act as their own publishers, supporting all the journals edited from within their institutions in a single place. Some of these portals are substantial in size. For example, the portals of the Universidad Nacional Autónoma de México, the Universidade de São Paulo, and the Universidad de Chile all host over 100 journals. These portals, together with thousands of other journals from Latin America, are hosted with the PKP's OJS.[36] Some preliminary data on the number of journals being published with OJS is now available,[37] although little is known about these journals, including whether they are peer-reviewed or not.

3.5. SELECTION CRITERIA

The first three initiatives outlined in this chapter collect a significant portion of the available data regarding scholarly journals in Latin America: Latindex, SciELO, and RedALyc. They have been working in Latin America for over 10 years and, through their collective efforts, have significantly contributed to raising the editorial quality of journals published in the region. All three initiatives post a list of require-

35 Molloy (2005) and Alonso-Gamboa and Russell (2012) present annotated lists of some of these sources in the Latin American context.

36 http://pkp.sfu.ca/ojs

37 http://pkp.sfu.ca/ojs/ojs-stats; http://pkp.sfu.ca/ojs/ojs-map

ments that journals must meet before they can be included in any of the databases. As these initiatives have grown and consolidated, they have slowly made their way into the evaluation criteria of universities and national science councils.[38]

The legitimacy afforded by national and institutional evaluation systems allowed the major OA initiatives to become synonymous with high editorial standards. They were not incorporated into the systems of evaluation by sheer existence or size, but rather by imposing basic and clear standards around editorial quality. These standards have come, over time, to define the editorial characteristics of a quality scholarly journal in the Latin American context. The oldest of the three, Latindex, set out a list of criteria to be used for entering the Latindex Catalog as early as 1997. The list now contains over 30 criteria such as using peer review, having an ISSN, and displaying an editorial board. SciELO and RedALyC have both adopted similar criteria as the requirements for inclusion.

The three lists have converged, with 11 elements common to all three lists and eight common to at least two (Table 3.2). In effect, these three initiatives have created a standard set of editorial criteria that all academic journals can use as a model. These three sets of editorial criteria and the inclusion of one or more of the portals into national systems of evaluation appears as an emerging trend, signaling the characteristics of what is considered a "high quality" Latin American publication.

The criteria can be generally grouped into three categories: those aimed at increasing internationality, those aimed at ensuring quality, and those aimed at improving metadata/indexing. Of those criteria for improving quality, perhaps the most important is the need for peer review. Peer review is widely considered the cornerstone of editorial quality, but it is surprising how many journals in the region still lack a well-defined editorial workflow with peer review by those outside of the journal's own editorial board.[39] The OA initiatives have brought peer-review to the forefront of conversations between editors and those doing journal quality assessments by making this requirement explicit and through workshops on editorial practices.

Other criteria are aimed at organizing the journals' metadata and increasing their visibility through improved indexing. In this category

38 A significant portion of this section can be found Alperin (2011).

39 During our own experience conducting workshops in 11 Spanish-speaking Latin American countries, we anecdotally found that there is still a poor understanding of how peer review should be conducted.

are the requirements to list article metadata on each page, listing author names clearly, listing of editorial teams, and other journal information. This point, seemingly trivial, can improve a journal's visibility in remarkable ways.

These criteria, representing a regionally defined standard, form the basis for all the subsequent description and analysis shown here for Latin American journals. They delimit the journal characteristics considered in-scope for the analysis. Again, as we acknowledged above, and as the Latindex Directory clearly shows, there is an incredible array of publications that do not meet these criteria (and surely there are even more not captured in the Directory!). There are also many other forms of research dissemination (i.e., monographs, which remain a prevalent and important form in the region, especially in the arts, humanities, and even in the social sciences). We regretfully do not include those in this analysis, as this chapter focuses on indicators of scholarly articles and the venues around which they are published – namely academic journals and institutional repositories.

Table 3.2
Selection criteria of Latindex, SciELO, and RedALyC

Requirement	Latindex*	SciELO**	RedALyC***
mentions editorial team	✓	✓	✓
has a minimum number of original research articles	✓	✓	✓
author names and affiliations are clearly labeled	✓	✓	✓
adherence to publishing schedule	✓	✓	✓
displays a table of contents	✓	✓	✓
has and displays ISSN	✓	✓	✓
displays submission and acceptance dates	✓	✓	✓
displays abstracts in English[a]	✓	✓	✓
displays keywords in English[a]	✓	✓	✓
peer-review with at least two external reviewers	✓	✓	✓
indicates a citation style	✓	✓	✓
has and displays a publishing entity and location	✓	✓	✓
has existed for a minimum period of time		✓	✓
displays article metadata at the beginning of article	✓	✓	
has defined focus and scope or objectives	✓	✓	

Requirement	Latindex*	SciELO**	RedALyC***
forms part of any indexing system	✓	✓	
requests declaration of originality	✓	✓	
displays URL on homepage (electronic only)	✓		✓
has minimum percentage of external authors	✓		✓
has minimum percentage of external editorial board	✓		✓
provides access to archives	✓		✓
displays name of director/manager of journal	✓		
display publishing schedule	✓		
mentions editorial board	✓		
displays copyright policies		✓	
minimum publishing schedule			✓
publishes a minimum number of articles per year			✓

Source: Alperin, 2012

* This list is a combination of the requirements for print journals and electronic journals wishing to enter the Latindex Catalog. Some of the parameters are omitted from this list, while others were merged into a single line item. Details an be found at http://www.latindex.org/latindex/catalogo.html

** Journals wishing to enter RedALyC do not need to meet all of these criteria. Some criteria are strictly mandatory while others are specified on the basis of "x number from this sub-list." Many of these line items are a combination of multiple requirements. Details can be found at http://redalyc.uaemex.mx/redalyc/media/principal/auxHemeroteca/criterios.html

**** Each country is free to specify its own criteria for entering SciELO. This list was taken from the SciELO Chile site. While the criteria for other SciELO sites are similar, it can vary on specific points from the selection shown here. Details can be found at http://www.scielo.cl/sr_scielocl/ CriteriosObligatoriosScielo.pdf

[a] Latindex does not explicitly specify English as a second language.

NB1: For the purposes of simplicity, this lists only the major requirements specified by all three portals and purposely leaves out some minor requirements present in only one of the three lists.

NB2: The requirements have not been literally translated in order to provide the commonality between the three separately worded lists.

3.6. CHOOSING WHICH DATA

Getting a sense of the number of journals actually published in a region is much more difficult than one would imagine, despite the numerous databases available for precisely this purpose. Recent work by Carvalho Neto (2013) shows the degree of this complexity by attempting to arrive at a number of active Brazilian journals. The author analyzes 13 lists of journals from various sources (including those listed here) and discovers the overlaps between sources are often small, and that numerous databases contain errors and omissions. This type of diverging coverage, even for a single country with well-established tracking of publications, certainly lowers the confidence that any combination of lists results in complete coverage. Yet, Carvalho Neto (2013) is optimistic that by crossing various databases, "it is possible to establish the number of Brazilian OA scientific journals and con-

duct an analysis of national *leading journals* and their features" (n.p.; emphasis added). While tracking all periodical publications in Latin America may be impractical, a similar approach can be used for identifying all Latin American journals that meet the minimum editorial criteria established by Latindex, SciELO, and RedALyc (summarized in Table 3.2).

To analyze the current situation in Latin America, we look at the overlaps between the journal lists available from the data sources above. It should be noted that the datasets were compared and joined using ISSN and eISSN numbers, where present. In a non- systematic check of apparent errors in joins, Octavio Alonso Gamboa (General coordinator of Latindex) already identified several typographic errors, missing values, and other errors in several of the datasets. This type of imperfect data is an unfortunate reality, but one that can be greatly improved.

As shown in Figure 3.3, Figure 3.4, and Figure 3.5, the Latindex Catalog (the list of journals meeting editorial criteria) is the largest of the existing datasets and in most cases approximates a superset of the others.[40]

We suggest that the Latindex Catalog provides the most accurate portrayal of all Latin American scholarly journals meeting Latindex's minimum set of editorial criteria. As we can see from the following tables, the coverage by subject and country varies greatly (Latindex, SciELO, and RedALyC) (Table 3.3). Similarly, coverage by country also varies significantly by source (Table 3.4). In both cases, we see that neither SciELO nor RedALyC provide coverage in proportions similar to those of in the Latindex Catalog. That is, neither provides a representative sample of countries or subjects when compared to the Latindex Catalog.

---

40 Without any manual correction of errors, the Latindex Catalog does not include 4% of RedALyC, 10% of SciELO, 19% of DOAJ, and 28% of Ulrich's. The higher numbers for DOAJ and Ulrich's are not as alarming as they seem, given that this comparison is to the Latindex Catalog, which applies editorial standards not present in the DOAJ and Ulrich lists. When compared to the Latindex Directory (the all inclusive list of journals), Latindex only misses 11% of DOAJ and 15% of Ulrich (Figure 3.4). Only 15 of the 154 Latin American journals in the WoS are missing from the Latindex Catalog (and only 10 from the Directory). A manual check of these discrepancies shows that, in fact, most of the apparently missing entries from the Latindex Catalog and Directory are in fact present. However, data entry errors or missing ISSN numbers in one of the datasets make the matching non-exact. Steps have been taken to correct these entries in the Latindex Directory and it is estimated that in early 2014 the Latindex Directory will be missing no more than 5% of any of the other lists.

Figure 3.3
Overlaps between Latindex, SciELO, and RedALyC

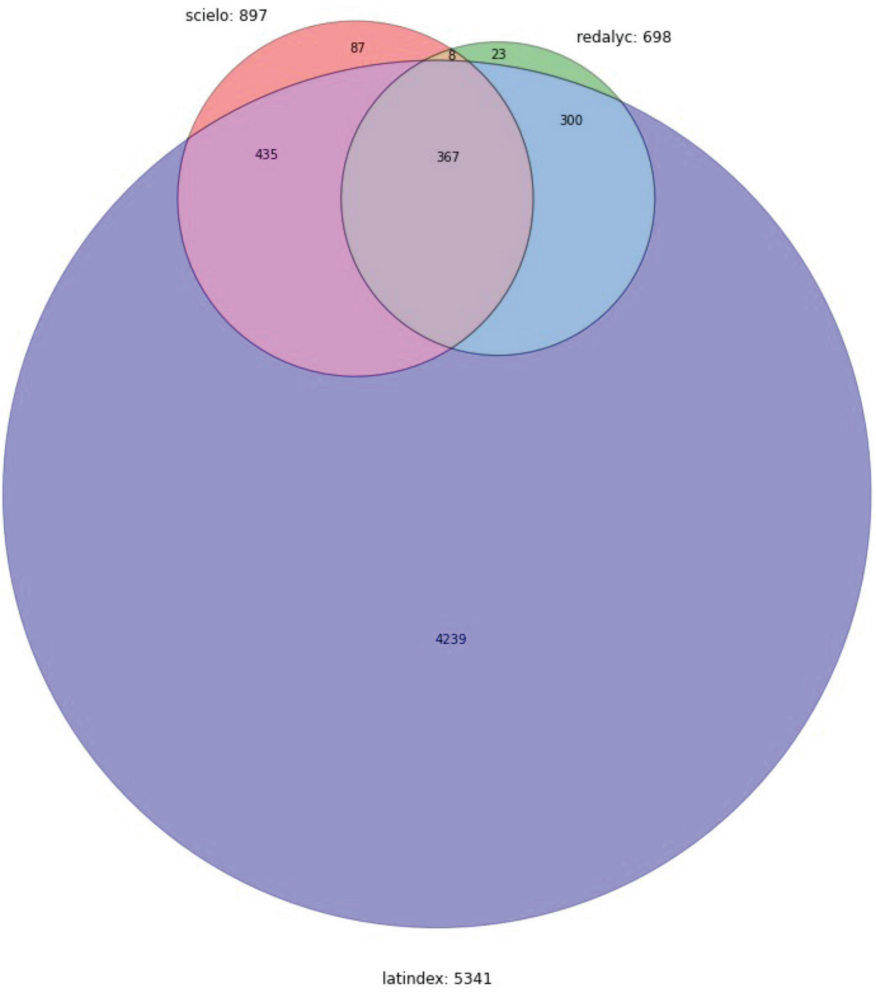


Figure 3.4
Overlaps between Latindex, DOAJ, and Ulrich

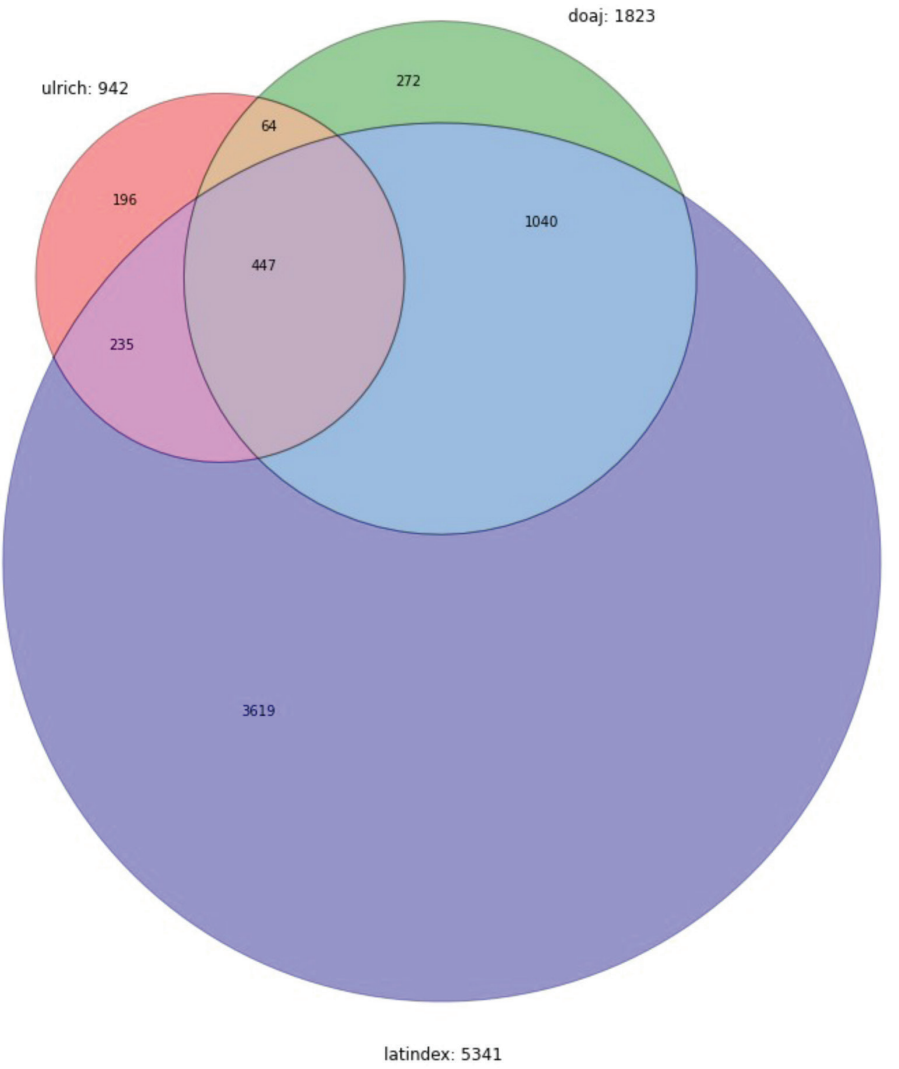


Figure 3.5
Overlaps between Latindex, WoS, and DOAJ

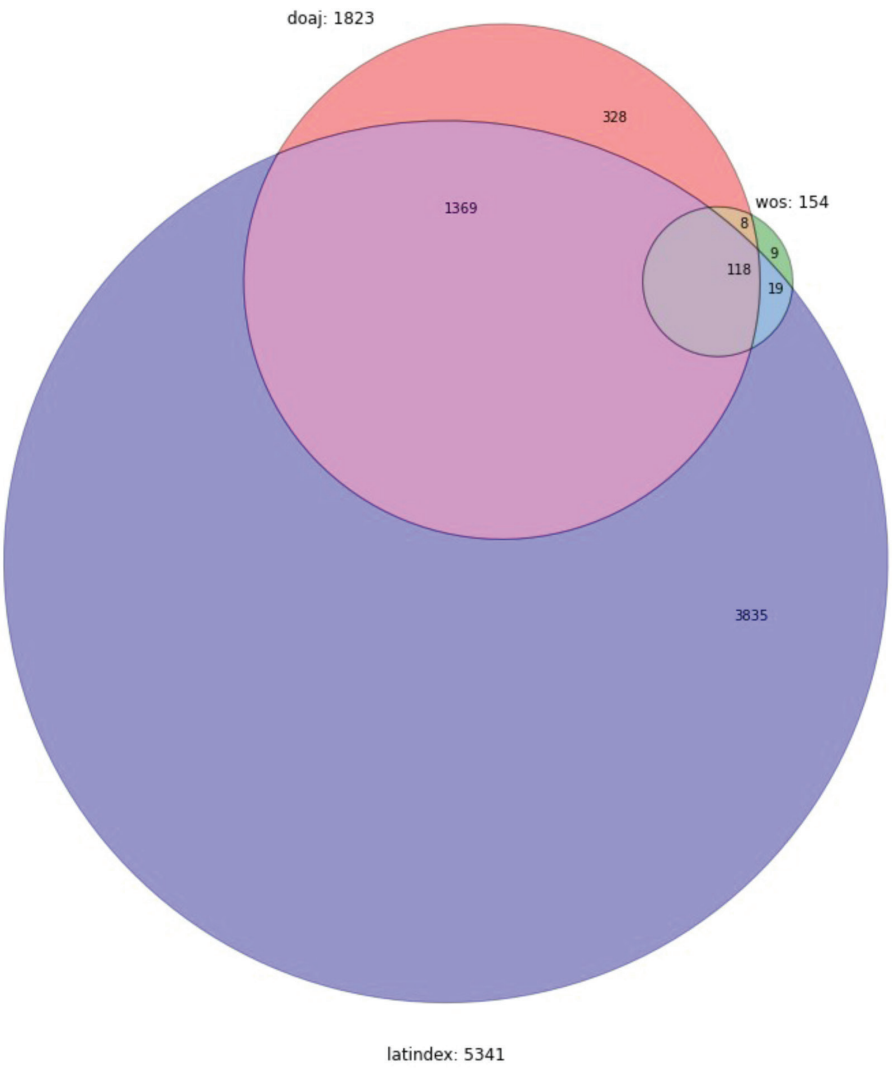


Table 3.3
Percentage and number of journals by subject area for Latindex, SciELO, and RedALyC

Subject	Latindex	Redalyc	SciELO
Unknown	-	4.30% (30)	10.48% (94)
Arts and Humanities	13.39% (724)	8.88% (62)	9.14% (82)
Agricultural Sciences	5.40% (292)	7.88% (55)	8.58% (77)
Medical Sciences	17.01% (920)	12.46% (87)	23.41% (210)
Natural Sciences	12.76% (690)	12.75% (89)	15.05% (135)
Social Sciences	50.24% (2717)	55.30% (386)	35.79% (321)
Engineering	6.82% (369)	8.74% (61)	6.58% (59)
Multidisciplinary	8.25% (446)	4.15% (29)	1.45% (13)

NB: Latindex allows for multiple subjects to be applied to each journal, so a journal may be counted more than once in the above table.

Table 3.4
Percentage and number of journals by country for Latindex, SciELO, and RedALyC

Country	Latindex	Redalyc	Scielo
Argentina	10.48% (567)	5.87% (41)	12.26% (110)
Barbados	0.02% (1)	-	-
Bolivia	0.46% (25)	-	-
Brazil	35.15% (1901)	22.06% (154)	34.11% (306)
Chile	6.29% (340)	9.74% (68)	11.15% (100)
Colombia	9.87% (534)	22.06% (154)	16.72% (150)
Costa Rica	2.07% (112)	2.15% (15)	1.67% (15)
Cuba	2.39% (129)	3.15% (22)	5.35% (48)
Ecuador	2.31% (125)	0.14% (1)	-
El Salvador	0.37% (20)	-	-
Guatemala	0.26% (14)	-	-
Haiti	0.02% (1)	-	-
Honduras	0.06% (3)	-	-
Jamaica	0.09% (5)	-	-
Mexico	14.92% (807)	24.50% (171)	13.04% (117)
Martinica	0.02% (1)	-	-
Nicaragua	0.63% (34)	-	-
Panama	0.39% (21)	-	-

Country	Latindex	Redalyc	Scielo
Paraguay	0.30% (16)	-	-
Peru	3.61% (195)	1.58% (11)	-
Puerto Rico	1.83% (99)	0.57% (4)	-
Dominican Republic	0.91% (49)	0.14% (1)	-
Uruguay	1.59% (86)	0.29% (2)	-
Venezuela	4.73% (256)	7.74% (54)	5.69% (51)

3.6.1. A note on data collection and processing

Journal lists from Latindex, RedALyC, SciELO, and DOAJ are all publicly available via the respective websites. The journal list from Ulrich's was obtained from the online version of the Ulrich's Knowledgebase using a subscription at Stanford University. The list of journals from the WoS was obtained through the Mimir Project at Stanford University. All lists are current as of mid–2013 and correspond to journals active in 2012 or 2013. Other data from SciELO, RedALyC, and CLACSO were obtained through the reports prepared as part of the project that funded the production of this chapter. All reports are available as appendices. In all cases, only data from Latin American, the Caribbean is included. Several of the initiatives also collect data on Spain and Portugal, or are international in scope, but journals from those countries is excluded from the analysis in order to focus on the situation in developing regions.

4. MEASURING GROWTH

Tracking the growth of OA is potentially less challenging in developing regions than for the world as a whole thanks to the regional initiatives that support and collect data about OA. Through these initiatives it is possible to garner information on the number of journals and articles published in the region. However, what is made clear in Figure 3.3 and Figures 3.4, none of the OA databases (e.g., SciELO, RedALyC, or DOAJ) are sufficiently representative of the publications from Latin America. Fortunately, regional OA initiatives in Latin America can work towards painting an accurate picture of the OA journal landscape.

The challenge of capturing the total output from the region, including all the articles that are published abroad and those deposited in institutional repositories, is a more daunting one. Institutional, national, and regional repositories are distributed unequally throughout the many higher education and research institutions, independent ini-

tiatives, councils, and governments, which makes an accurate count or even an approximation difficult to arrive at.

The following sections present what is currently known about the growth of OA articles in these developing regions, including those published in regional journals and those deposited in institutional repositories.

4.1. PUBLICATIONS

When looking at the changes over time through the lens of a single data source, we observe erratic patterns from year to year. However, they likely reflect the immediate capacity at the data collection point rather than the real growth taking place. Figure 4.1 shows the growth rates of several of the already described data sources. Note that Latindex and DOAJ both collect the start date of the journal (date of first publication), while SciELO and RedALyC collect only the insertion date (i.e., the earliest issue published in the database). We notice very erratic growth rates in the early years of both SciELO and RedALyC as the initiatives got off the ground, but neither has yet to achieve stable growth (in both cases the graphs are shown after the first two years once growth began to stabilize). The RedALyC start dates shown were inferred from the RedALyC insertion date, and do not necessarily reflect the actual start date of the journals. By contrast, the Latindex Catalog shows that when using the start date, there is stable growth of between 5% and 6% annually – even in the years that 25% new entries were being added to the Catalog. While the Latindex Catalog gives a sense of the growth of publications, there is a challenge in disentangling the growth of journal publishing in general from the growth of OA publishing.

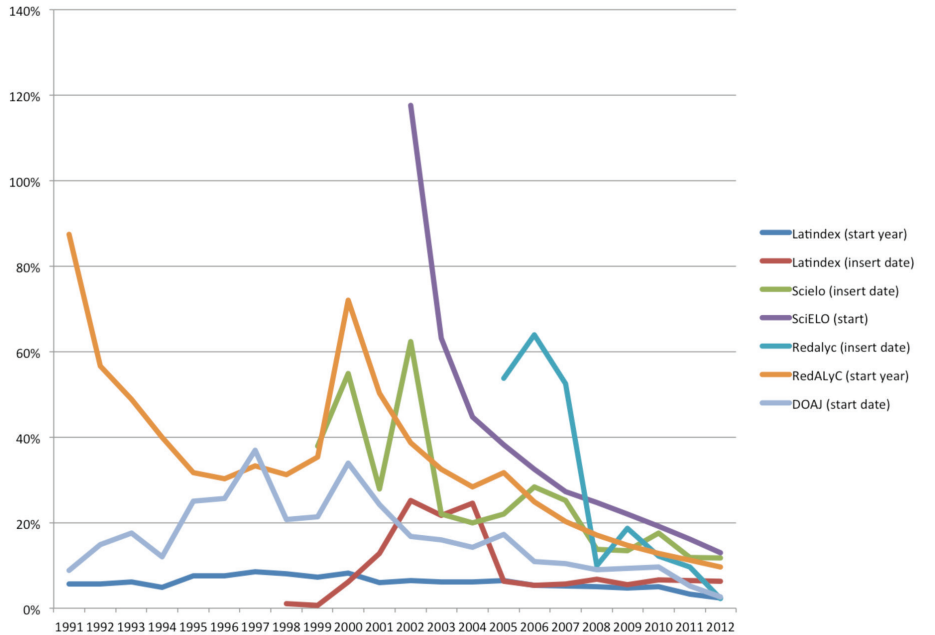
It is possible (even likely) that, as the figure above shows, the growth rate of OA journals is higher than that of journals on the whole. However, we must caution strongly about drawing any specific conclusions from a single data source.[41] The goal of this section of the report is to disentangle the growth of the available datasets from the growth of OA itself.

The inclusion of the first publication date field in the DOAJ database helps to address the database growth problem, although it re-

41 This caution extends not only to the growth of publishing activities, but also to any study that uses any individual dataset to draw conclusions. For example, scholars have recognized that the inclusion of new journals can have a measurable effect on the structure of citations at the subfield, field, and global levels (see Leydesdorff & Cozzens, 1993; Leydesdorff et al., 1994; Michels & Schmoch, 2012). Such distortions are especially egregious when the dataset being used is not entirely representative of the field being studied.

mains unknown how representative the journals included in the DOAJ are of all OA journals. In cases where a first publication date is not available, an approach similar to that used by Michels and Schmoch (2012) can be used. The authors estimated first publication dates by looking at the volume numbers to determine the date of the first article. They found that when looking only at growth due to new journals (as opposed to those added to the database) there is a marked difference in the estimated growth rates, and that this varied wildly depending on the process leading to additions in the database.

Figure 4.1
Journal Growth as Percentage of Previous Year

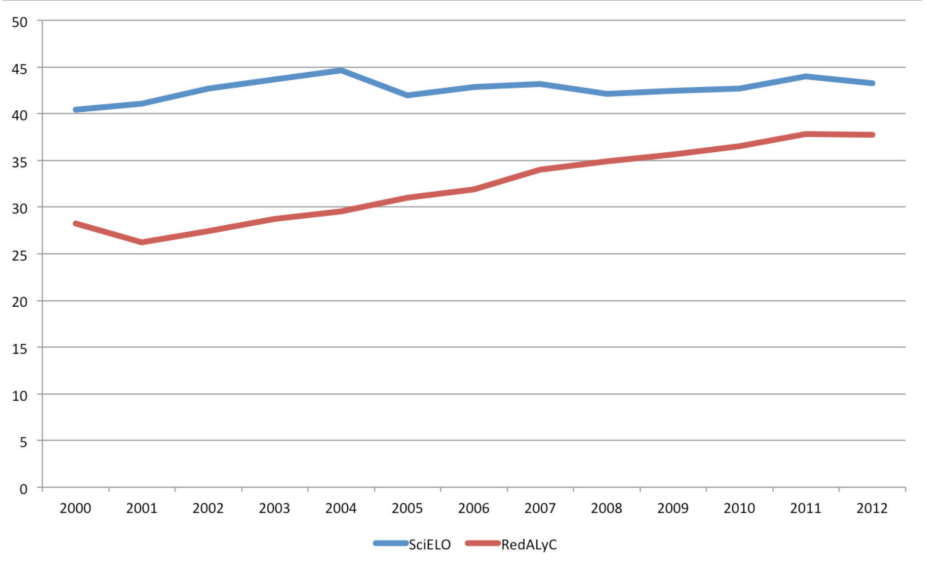


Another approach to assessing the growth of OA has been to use survey methods. The most comprehensive of such surveys have been those by Laakso, Björk, and their colleagues (e.g., Laakso & Björk, 2012; Laakso et al., 2011). Although the authors rely heavily on the DOAJ to identify OA journals, they also conducted surveys in an attempt to identify other journals not listed. More importantly, by col-

lecting information at the article level instead of relying only on journal metadata, the authors uncover important characteristics about OA publishing. Notably, they discover that the growth rate of OA articles is higher than that of journals (i.e., journals are publishing more articles per year than before). This suggests a shift away from the previously held notion that OA journals published few articles per year, at least in comparison to subscriptions journals (McVeigh, 2004).

The extent to which this holds in developing regions is still unknown, although it can be examined by using the article-level data available from both SciELO and RedALyC. On SciELO, article volume has remained relatively constant between 2000 and 2012, fluctuating only between an average of 40 and 45 articles per journal per year (Figure 4.2). However, on RedALyC the average has increased from 28 to 38 articles per journal per year.

Figure 4.2
Article Growth as Percentage of Previous Year



Neither dataset can confirm the hypothesis that OA journals are publishing on average more articles per year than before, but the difference highlights the importance of not drawing conclusions from a single – and potentially non-representative – source.

4.2. REPOSITORIES

There has been abundant support for OA repositories in Latin America. Peru and Argentina are the only two countries that have mandated self-archiving for all federally funded work, in both cases with provisions for support for institutional repositories. Brazil has had a similar proposal presented to congress, but it has not yet been passed into law. Since 2012, the number of OA mandates has grown from 17 to 24 in Latin America according to the Registry of Open Access Repositories Mandatory Archiving Policies (ROARMAP).[42]

To accompany these mandates, there are an increasing number of institutional repositories. In the Registry of Open Access Repositories (ROAR),[43] Latin America currently has 329 registered repositories, up from 262 in 2012 and 227 in 2011 – a 44% increase over this two-year period. A significant number of these repositories are indexed in Directory for Open Access Repositories (OpenDOAR),[44] which provides several statistics on these repositories, including content types, languages, and the operational status. However, OpenDOAR does not index the bibliographic records, so it is unknown how many records are available and how many of these provide access to the full-text.

This limitation may be corrected by regional efforts underway in Latin America to integrate the region's institutional repositories. The member countries of the Latin American Cooperation of Advanced Research and Education Networks (RedCLARA)[45] have begun the Network of Federated Institutional Repositories of Scientific Publications (known as LA Referencia).[46] As agreements are reached that guarantee interoperability between repositories, including standards for how the bibliographic records are collected, it can be expected that more accurate information on the contents of institutional repositories from the region will be available. Through cooperation between LA Referencia and OpenDOAR, we can also expect more comprehensive coverage of existing repositories from the region.

To measure the growth of OA from regions such as Latin America, initiatives like OpenDOAR can prove invaluable, but they remain limited as long as they only provide statistics on the repositories themselves. Projects like LA Referencia are a step in the

42 http://roarmap.eprints.org/

43 http://roar.eprints.org

44 http://www.opendoar.org/

45 http://www.redclara.net/

46 http://lareferencia.redclara.net

right direction, as repository interoperability would facilitate the collection of article-level statistics on a wide scale. Some repositories, such as CLACSO's Virtual Library, already provide statistics on the growth of their collections,[47] as surely other repositories do, but these remain scattered and difficult to aggregate in order to form a more complete picture of the self-archiving activities of any region.

5. MEASURING REACH

We find it useful here to highlight two important distinctions related to reach. First, the distinction between ***reach*** and ***impact*** of research borne out in the definitions provided at the outset of this chapter. In that definition, we clarify that by reach we are referring to both the demographics and characteristics of the readership of scholarly content and to the characteristics of the channels in which the research is disseminated. Under our definition, download counts are a measure of reach, but citation counts are not (they fall under our definition of impact). The second distinction is that between the article "views" (i.e, HTML and PDF downloads) and all other measures of reach (i.e., social media mentions, social bookmarking).

Little is known about the extent to which research published in developing regions is circulated and read. This is true of readership generally, not just in developing regions, as publishers do not generally disclose detailed usage statistics.[48] However, for developing regions details on the demographics of the readership is of critical importance, as governments and funding agencies strive to focus their limited resources in serving their constituents.

There is a sense that developing regions invest in science in the hope that it results in development in a broadly defined sense. Public investment in R&D, including the significant investment necessary for publishing scholarly journals free of charge for both authors and readers, which is largely the case in developing regions, is justifiable only if it contributes to the public good. The creation of a set of indicators that capture the reach of OA research from developing regions is therefore a valuable policy tool.

Indicators of reach are also immediately useful for authors, editors, and institutions, not just for evaluation purposes, but also to help justify the significance, relevance, and value of their work to complement citation information. And, even in the absence of scholarly

47 http://biblioteca.clacso.edu.ar/estadisticas/crecimiento/

48 The obvious exception here is PLOS (http://www.plos.org), which makes all the metrics they collect available for download at the article level.

citations, there are many ways in which a research article can contribute to the public good without ever receiving citations.[49]

5.1. DOCUMENT VIEWS

We separate document views from other forms of reach because they represent the lowest level of engagement that is possible with an article (Lin & Fenner, 2013). View metrics are also unique in that they are potentially available for every OA publication since, by definition, an OA article must be online. In fact, most OA journals today already track users to some degree, whether through dedicated systems or through standard web analytic tools (i.e., Google Analytics). View statistics also present a unique set of challenges that must be addressed if they are to be used as indicators of the reach of OA.

The same thing that makes usage statistics so pervasive is what makes them problematic as indicators of reach. Usage statistics are simple for journals to collect themselves, which naturally leads to a wide range of implementations that make them incomparable. Despite existing standards such as COUNTER (COUNTER, 2012), which aims to standardize implementation, the journal interfaces affects which content users reach which articles – one again making reports incomparable (Davis & Price, 2006). More recently, COUNTER has extended its code of practice, which operates at the journal level, to the article level with the PIRUS code of practice (Shepherd, 2011). PIRUS suffers from many of the same difficulties of COUNTER with the additional difficulty in its implementation in developing countries due to its reliance on Digital Object Identifiers (DOI), which are not as prevalent in the global South (Alperin, 2013a).[50]

---

49 A research article that is used, for example, for didactic purposes in an undergraduate classroom contributes to the development of human capital and to the strengthening of a higher education system. A policy recommendation in a paper being taken up by a government agency can change the life of the citizenry. Information about new medical treatments in the life of a someone suffering from a disease can help lead the patient to better manage their illness, or to simply to be given hope about their prognosis. A study of the effects of an intervention can help an NGO to adjust their programs to better serve their community. The list could go on and on, but the point is that citations only capture only a particular type of usage and therefore only give a glimpse into the reach of any article.

50 Projects like MESUR (Bollen et al., 2007) have successfully integrated usage data from various sources, but only by investing "significant energy" (Bollen et al., 2008, p. 231). Although the lessons and examples from MESUR are promising, they are difficult to replicate. Moreover, the project itself did not produce a publicly available dataset, nor did it include a significant number of OA publications or those from developing regions.
More recently, the Open Access Statistics (OAS) project (http://www.dini.de/projekte/oa- statistik/english/) has successfully implemented a centralized infrastructure to

Despite the difficulties in aggregating and comparing document views across publishers, download statistics are still a valuable resource in understanding the nature of OA in developing regions. Although in their current non-standardized form they should be used with caution, especially for evaluation purposes, they provide a first glimpse at the level of knowledge consumption, which some claim is largely mismatched to the level of knowledge production (Velho, 2004).

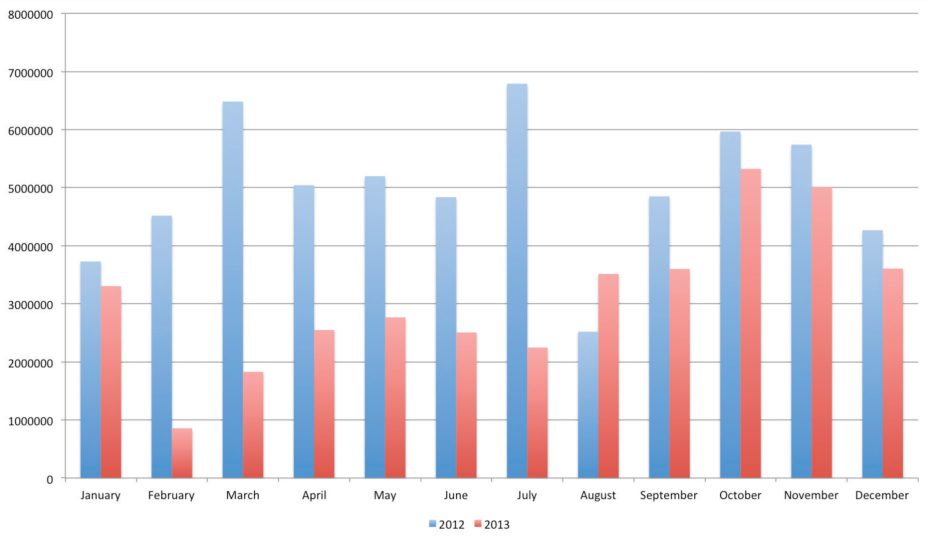
Both RedALyC and SciELO provide some usage statistics to the public. In the case of SciELO, they are available from their website, but only for the Chile, Brazil, and Public Health collections (as well as other collections in the following tables, replicated from the attached report produced by RedALyC) one can get a glimpse at the characteristics of the usage of Latin American research. The RedALyC download data, some of which is shown in more detail below, provides us with a sense of what can currently be inferred from usage statistics in general, but also shows us some of the difficulties in comparing results to one another and the challenges of arriving at an accurate count of users – even using data from a single publisher.

For example, Figure 5.1 below shows a drastic change in the number of downloads per month between 2012 and 2013 on the RedALyC website. A noticeable drop (of over 40% in all cases) can be seen in the months following February – at the same time that RedALyC overhauled the web portal to include better identification of "robots."

---

collect and process usage data from OA publishers and repositories in Germany. The OAS experience agrees with the PIRUS code of practice in that a central clearing house is necessary to achieve the type of integration called for here (Herb, 2010). In Latin America, a less ambitious effort has also began through the *Quality in the Open Scholarly Communication of Latin America* project (http://flacso.org.br/oa/calidad-en-la-comunicacion-cientifica-abierta-de-america-latina), which aims to aggregate download statistics from SciELO, RedALyC, and journals using the PKP's OJS. It operates through a central service that acts as a clearinghouse for statistics from the above-mentioned sources. It is expected to be in operation in early 2014. Details on the service can be found in the online documentation at http://scielo.readthedocs.org/projects/ratchet/en/latest/

Figure 5.1
PDF Downloads from RedALyC



So while comparisons between years and between publishers can be problematic, there is still some highly relevant information contained in the RedALyC and SciELO usage statistics. For example, the overall pattern of traffic shown in Figure 5.1 is generally consistent from year to year (July 2012 appears to be an outlier in this case), and we see a decrease in traffic over the holidays, bottoming out in February and another dip in the summer months. This is similar to what has been reported for SciELO Brazil, although SciELO's traffic bottoms out in December and January and is back on the rise again by February. SciELO Brazil also sees a significant dip in downloads in July (coinciding with the Southern hemisphere's winter break). These patterns are both consistent with academic use (i.e., rising and falling with the academic calendar), although no firm conclusions can be drawn.[51]

Not only are the number and time of accesses important, but also the location of the readers. While SciELO does not (yet) make country of access available, RedALyC does track downloads by country and

51 Similar speculations has been made about the SciELO Brazil usage data (Adams, 2013).

month.[52] These types of patterns provide a global sense of who reads Latin American research, with the caveat that access to sites like SciELO and RedALyC may not be representative of usage of all OA journals. If these data were made available at the article level, along with the article metadata, it would be possible to draw further inferences about the type of content that receives usage and where, leading to a clearer sense of the audience that accesses OA journals.

5.2. ALTMETRICS

Thanks to altmetrics, there are now other indicators of reach beyond document views. As we will see, most of the available metrics imply a level of engagement that would fit under our definition of impact.[53] So while a specific mention (i.e., in social media, a bookmark, or a blog post) is an indication of use, it also provides information on the extent of the reach and provides evidence of one of the channels in which the research is being disseminated. They tell us the networks the article is being circulated on, which may in turn be linked with information about the audience being exposed to the research.

Taking the popular social media source Twitter, for example, tweets are an indication that a given article is circulating in that particular social network, which provides us with evidence of the way in which people share and disseminate research. There is a lot of potentially useful information available if we also look at the network of Twitter users. Topsy,[54] a popular service that provides Twitter analytics, distinguishes between tweets from average users from those from "influential" users (those users that have a lot of followers). ImpactStory also offers such a service. Such measures can give a sense of how many people were exposed to a given article (i.e., how many people it reached), even if those people never clicked through to view the article. Of course, Topsy and other services like it (i.e., Netbase, Crimson Hexagon, Brandwatch, Simply Measured),[55] are able to calculate a more precise count of the number of users who have potentially seen a tweet. The service ThunderClap,[56] which can be used to promote a

52 These numbers are not publicly available, but RedALyC intends to make them public.

53 See (Lin & Fenner, 2013) for a proposed nomenclature for classifying sources that uses the various levels of engagement (from viewed, to saved, to discussed, and finally cited).

54 http://about.topsy.com

55 http://www.netbase.com; http://www.crimsonhexagon.com; http://www.brandwatch.com; http://simplymeasured.com

56 https://www.thunderclap.it

cause, calls this exposure "social reach," a useful concept in the context of OA indicators. Sources like Twitter could provide information on the geographic location of those people who have been "reached," although one study has shown that only "8.2 percent of all users [during the studied period] had either Place or Exact Location information available for their tweets" (Leetaru et al., 2013, n.p.). Although only some tweets have geolocation encoding, Topsy claims that it can infer the country of origin of 95% of tweets and the city of origin of about 25% (MacMillan & Wakabayashi, 2013). Currently, only Altmetric.com offers this "upper bound" of followers and geographic location of those who tweeted, although it does so with limited accuracy). It should also be noted that all of the services that provide tweet analytics are commercial entities that provide their services for a fee.

A similar type of analysis could be done on other social networks, such as Facebook and Google+, although there are some limitations as much of the activities on these sites are not publicly available (the way that all Twitter data is). While the potential is there to track the "social reach" of links posted on Facebook, the current API only provides information on the number and types of mentions ("shares" and "likes") and a count on the number of users that mentioned a link. Facebook does collect location information for its users and so that data is potentially available, although not currently exposed (even in aggregate form), through its standard API.

There are some limited efforts (notably, again, by Altmetric.com) to also capture articles in news sources. These types of mentions represent a very different reader demographic than those who typically read research articles. Even if not in its original form (likely interpreted by a reporter), citations by news outlets have the potential to capture a reach of OA that is never seen on the article page, and thus missed by download counts. This is an underexplored area, especially in the developing region context, but one that is currently being researched (Marin & Petralia, 2013).

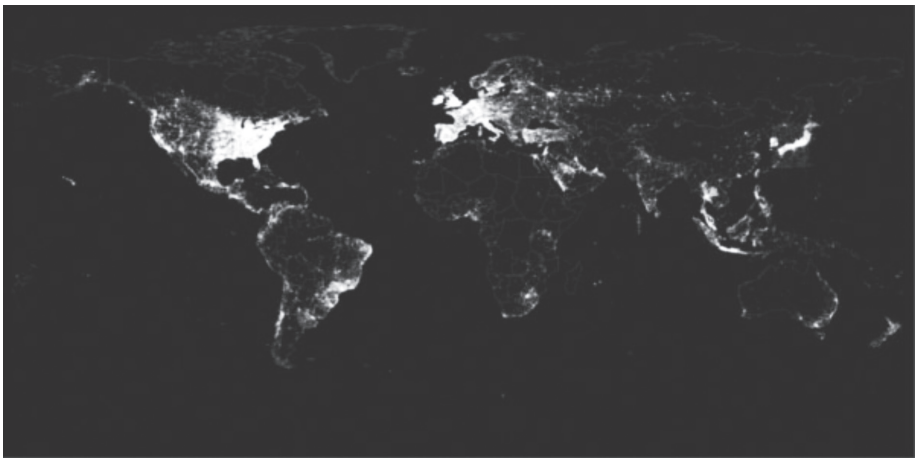
Finally, mentions on websites or blogs present yet another potential way to capture the extent of the reach of an article. Blogs and other online sites could provide secondary view statistics, including geographic location, of the pages that mention an article. None of the current providers do this at the moment, although ImpactStory has begun offering altmetrics on specific blog posts.[57]

All the altmetric sources, be they social networking sites, blogs, or other services, present a new channel of communication that is not captured by looking at direct views or downloads of research from journals

57 http://blog.impactstory.org/2013/12/16/blogs

or from institutional repositories. In most of the cases it is possible, at least in theory, to capture not only the number of mentions (as is currently done by most providers) but also to capture information about the audience who is exposed to the altmetric source. This information could include the size of the audience, the number of specific views, and even geographic location. However, even if the information could be captured, the level of penetration of the altmetric sources in developing countries should not be ignored (Alperin, 2013a, 2013b). In particular, one must be cautious when interpreting the demographics gathered from the altmetric sources, given that the usage of these sources is not evenly distributed around the world. Figure 5.2 illustrates this point by showing the representation of all tweets with location data from a one month period, indicating that Twitter does have relatively broad coverage (Leetaru et al., 2013), although some countries and entire regions do not appear to be well-represented (such as Bolivia, Perú, and Cuba).

Figure 5.2
Twitter Coverage (Leetaru, Wang, Cao, et al., 2013)



6. MEASURING IMPACT

The impact of OA has been under great scrutiny for over a decade – almost since the inception of the OA movement. Many studies have examined the links between OA, downloads, and citations[58] (e.g., An-

58 So many, in fact, that an OA impact bibliography was set up: http://opcit.eprints.org/oacitation-biblio.html

telman, 2004; Brody & Harnad, 2004; Eysenbach, 2006; Harnad et al., 2004; Lawrence, 2001). These studies have largely focused on the impact of OA as measured through what is known as the "citation advantage of OA," which is only natural given that citation metrics have been the defacto metric used for the evaluation of impact (in particular through Thomson-Reuter's Impact Factor).

Much effort has been spent on trying to measure the "impact" of science generally –primarily its economic impact and its effect on innovation (Castells, 2011; Maxwell & Stone, 2013). Naturally, national systems of innovation (e.g., Lundvall, 1988; Nelson, 1993) in particular have been the primary focus of research due to their policy implications (Irvine & Martin, 1984; Leydesdorff & Meyer, 2006; Nelson et al., 1982). However, all too often the approach – as it pertains to scientific publications – has been limited to studying the citation links between scholarly documents and patents (for an overview, see Smith, 1998). This approach overlooks the many other ways in which the scholarly literature contributes to development.

In this chapter, however, we use a much broader definition of impact, one that includes citations but does not limit itself to them. By looking at a broad range of indicators of usage, it will be possible to gain a sense of the places and people that Latin American research affects. However, we acknowledge the importance of impact within the academic realm as measured by citations. So again, as with document views, we draw a distinction between citation-based indicators and the alternatives provided by altmetrics in the following two sections.

6.1. CITATIONS

Before the creation of SciELO, the only sources of citations of journals from developing regions were the WoS and Scopus – two databases that underrepresent the developing world. Putting aside issues of equity, the underrepresentation and shear low number of journals from developing countries means that journals that are geared towards the developing world will have less of their citations counted than other journals in their dataset. It is known, however, both by looking at the WoS (Collazo-Reyes, 2013; Collazo-Reyes et al., 2008) and at SciELO Brazil (Meneghini et al., 2006) that journals do precisely this. While some attract international citations, others appear to be geared towards local or regional audiences. As a direct consequence, the citations in these journals are simply not captured in the dataset and their resulting IFs are kept artificially lower. As a consequence, the majority of Latin American journals have historically had their IFs in the fourth quartile (Luna-Morales & Collazo-Reyes, 2007; Packer & Meneghini, 2007).

The same problem exists when looking at the citation counts calculated by SciELO (the difference, of course, is that SciELO is only capturing local and regional citations and omitting all those from outside the region). The partnership between SciELO and Thomson-Reuters aims to address this challenge by combining citations from the WoS and SciELO, but the data is not yet publicly available. Even so, this solution is not a panacea. Both databases, even when combined, present a series of problems and limitations that must be considered when trying to understand to the citation impact of a developing region such as Latin America.

The biggest challenge for determining the impact of journals, and articles, from developing regions is that different coverage means different citation counts. Figures 6.1 and 6.2 present the average IFs of journals from various Latin American countries, as calculated by Thomson-Reuters' WoS and SciELO respectively, in the period following the year 2000. Both paint very different pictures of the progression of citation impact over time at the country level. While the journals in the WoS appear to maintain relatively steady (except for Brazil), the SciELO IFs show an increase for all countries. On the surface, these differences could be explained by the understanding that the journals included in SciELO are receiving more citations over time from local or regional publications, whereas the ones in the WoS are receiving similar number of citations from their international counterparts (Figure 6.1).

Note, however, that the differences are not only caused by the differences in coverage of the journals receiving the citations (those whose IFs are being graphed). We still observe differences when plotting the average IF of the same set of journals for both databases (those journals indexed in both) (Figure 6.3). In these cases, the discrepancy in average IF is caused by the difference in the citing journals instead. Interestingly, we observe some similarities in the fluctuations of the IF,[59] even as there is a persistent gap in the IF values between SciELO and Thomson-Reuters. The graphs in Figure 6.3 also emphasizes the importance of examining local and regional citations, as the SciELO IF, which captures only Latin American citations, is higher in all cases. Unfortunately, these figures still give an unsatisfactory explanation of the impact of Latin American journals at the country level, even as they serve to highlight some of the difficulties in drawing conclusions at the field and subfield levels.

---

59 A more careful analysis shows that except for Brazil and Mexico, the trend lines for the period following 2004 do not have slopes that are significantly different from one another.

Figure 6.1
Average Impact Factor by Country of Latin American Journals Since 2000 According to WoS

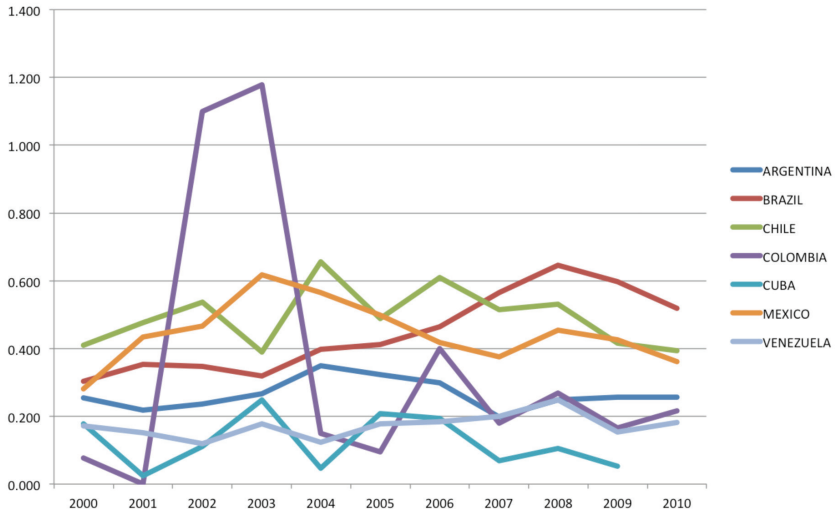


Figure 6.2
Average Impact Factor by Country of Latin American Journals Since 2000 According to SciELO

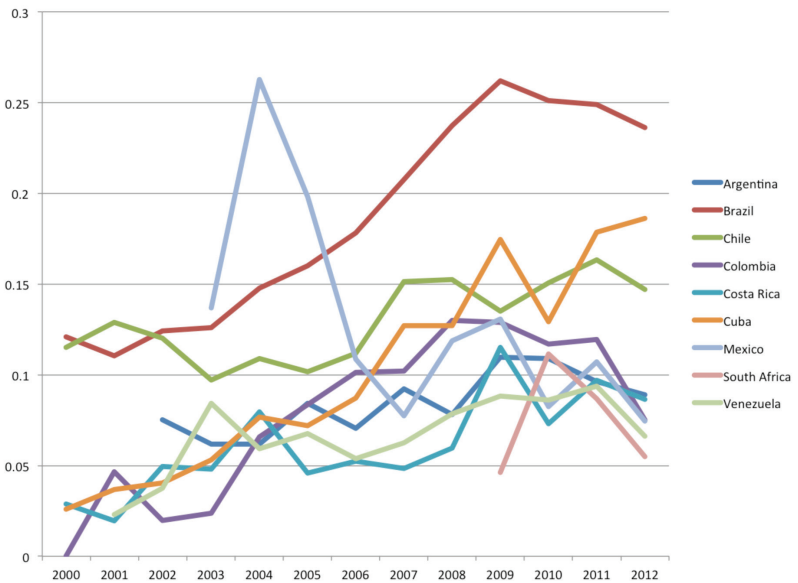
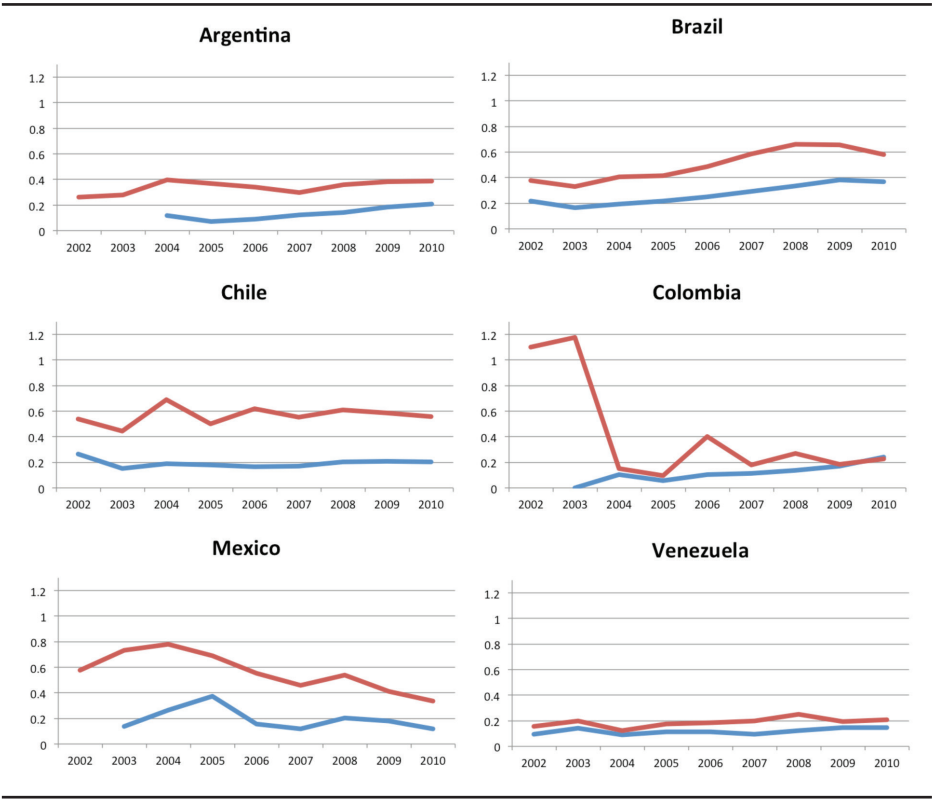


Figure 6.3
Average Impact Factors for Journals in both SciELO and WoS



The issue of database coverage is primary among these. The effects of the growth of the datasets that, at least in the case of WoS, "is mainly related to a change in the editorial policy . . . rather than to a change in the [Latin American] scientific community" (Collazo-Reyes, 2013, p. 207) itself has important effects on the resulting citation counts. Unfortunately, this issue is unavoidable without access to an all-encompassing set of articles and their citations.[60] The only remaining option, if citation impact is of interest, is to be cautious when interpreting the results of any analysis.[61]

60 The closest thing to an all-encompassing dataset is, of course, Google Scholar. Unfortunately, Google Scholar data cannot be mined and, even if it could, it comes with a whole set of other problems (see Jacsó, 2012).

61 It warrants mentioning that there is a growing body of literature on what is known as the "OA citation advantage" – the supposed increase in citations of articles

6.2. ALTMETRICS

We know that the presence of scholarly articles in online channels of communication is growing, but it is still unknown to what degree altmetrics can be used for capturing impact. There is a definite sense that the numbers provided by altmetrics in some way indicate attention, influence, or even impact, but it is still unknown how. There are, as of yet, few large-scale studies and most altmetrics research ends with a call for further study of the reliability, validity, and context of the available metrics (Haustein & Peters, 2013; Liu & Adie, 2013; Wouters & Costas, 2012).

Most altmetrics studies so far have focused on the relationship between altmetrics and citation metrics. So far, the strongest correlation between altmetrics and citations were found for F1000 Prime recommendations (Li & Thelwall, 2012). However, several studies have shown a moderate level of correlation between saves in the bibliographic manager Mendeley and citations (Bar-Ilan, 2012; Bar-Ilan et al., 2012; Li et al., 2012; Priem et al., 2012). Mentions in blogs have also shown to have a positive correlation with citations (Costas et al., 2014; Fausto et al., 2012), although this appears to be heavily influenced by the makeup of bloggers themselves and the journals they tend to blog about, both of which introduce a strong bias for high-impact life science journals (Shema et al., 2012). Twitter, the most prevalent of the social media altmetrics sources (Thelwall et al., 2013), was found to have a low (but positive) correlation to citations (Eysenbach, 2011; Haustein & Peters, 2013). Eysenbach (2011) also found that highly tweeted papers could also be used as early predictors of citations, something supported by other scholars (e.g., Shuai et al., 2012; Thelwall et al., 2013). In short, there appears to be a low to moderate correlation between altmetrics and citations, at least in the specific journals and disciplines studied (primarily well-known English-language natural and life science journals).[62]

However, most altmetrics and usage studies also express that these new metrics capture a different "dimension," "flavor," or "type"

---

that are OA versus those that are in subscription journals. For readers interested in this phenomenon and the studies supporting and contradicting the effects, see the following bibliography, especially the "rapid reader" section (Hitchcock, 2013). Note, however, that the effects are not of concern in the present study, as the subscription model is virtually non-existent in Latin America. This study is concerned with documenting the impact of OA publications without using toll-access publications as a counterfactual.

62 Haustein and Peters (2013) and Torres-Salinas et al. (2013) also summarize many of the studies mentioned here and provide some details on sample sizes and correlation coefficients (omitted here for brevity).

of impact than citations (Bollen & de Sompel, 2008; Bollen et al., 2009; Costas et al., 2014; Eysenbach, 2011; Haustein & Peters, 2013; Torres-Salinas et al., 2013). Not only are altmetrics and usage metrics different from citations, they are also different from each other (Bollen et al., 2009; Thelwall et al., 2013). The implication is that altmetrics "should not be considered as alternatives to citation-based indicators, but rather as complementary" (Haustein & Peters, 2013, pp. 18–19).

Unfortunately, no work has been done to explore the impact of altmetrics in developing regions. So, while the studies above provide a sense of the type of impact measures that might be uncovered from altmetrics, it is an empirical question whether these effects and differences can be replicated by using altmetrics on scholarly articles from developing regions.

7. RECOMMENDATIONS

The developing world has proven to be a leader in OA worldwide, especially Latin America. Much of the world benefits from many of the practices surrounding the production, collection, and dissemination of scholarly content from the regional initiatives in Latin America. The OA initiatives described in this chapter offer sustainable alternatives and should be seen as emerging best-practices for developing and developed regions of the world alike.

This chapter has covered in detail the sources of available data regarding research and scholarly outputs, with a focus on OA initiatives from Latin America. The following sections present those practices that we believe should be continued, expanded, or implemented to improve the quality and quantity of available indicators. We offer recommendations on improving the available data, as well as for improving the ability to track growth, impact, and reach of OA in developing regions.

7.1. AVAILABLE DATA

Most of the effort by current OA initiatives has been focused on the improvement of metadata on everything that is published in these regions. As was mentioned in the previous section, tracking the scholarly output is essential for putting other indicators in context. However, it is not enough to know the how much is being produced, where it originated, and at what pace. The level of uptake of this research could potentially serve as a marker of the research's usefulness, its audience, its potential impact, or even equality of the work. Yet, as this chapter has highlighted, very little is known about the reach of research from developing regions, even as so much emphasis is placed on its production.

To improve the quality of indicators on reach, we therefore recommend the following:

1. *Use standards such as COUNTER and PIRUS* for tracking downloads and document views. Even though there will continue to be difficulties in comparing document view data across providers, the compliance with standards will make this task simpler and more reliable.
2. *Track use at the article level, not website or journal levels* so that more detailed information can be inferred from the article's usage. While data at the journal level makes it possible to detect patterns in access by country of publication, language, and discipline, data at the article level makes it possible to track individuals (authors) and specific topics.
3. *Use unique identifiers and include them in the URL* to simplify tracking a document across the web. When the tracking of articles was only done through citations, it was possible to rely on the documents metadata. However, as non-academic sources become of interest (i.e., Twitter and Facebook), it is necessary to have a simple and unambiguous way of identifying when a document is mentioned. Unique identifiers, such as DOIs serve this purpose.
4. *Ensure identifiers are propagated across all versions of a document* so that statistics about multiple versions (i.e., a post-print and a published version) are counted together.
5. *Capture more than counts from altmetrics sources* wherever possible. The PLOS ALM application, for example, tracks the time when the data was collected to supplement sources that do not provide time stamped events.
6. *Conduct surveys of users* from websites and portals that provide access to content. Demographic surveys can provide a better sense of *who* the audience is, while download and altmetrics data can capture the number and location of those users.
7. *Share data used to produce indicators* as openly and widely as possible. Any collected data regarding scholarly outputs should be minimally displayed to the end-user, but preferably available to download en masse in machine-readable forms. Ideally, all data could be programmatically queried and manipulated, and be free of restrictions for use (i.e., distributed with licenses that permit data mining).

7.2. GROWTH

Tracking the growth of scholarly output generally, and OA outputs in particular, is fundamental to understanding other aspects of OA. The growth of OA must be tracked against the growth of publishing if it is to provide an indication of the success of OA models. Similarly, the impact and reach of OA only make sense in as far as it can be differentiated from the impact and reach of scholarship more broadly. This is of special interest to those promoting and supporting OA, especially, as is the case in Latin America, to the governments that are using resources to ensure much of the region's output is freely accessible.

To improve the tracking of the growth of OA over time in developing regions, we recommend the following:

1. *Identify OA publications in all journal and databases*, preferably using standard metadata tags. The Latindex model for tracking publications has been successful (and is recommended for other regions), but it should be expanded to track the journal's business models. To this effect, it is possible the recently launched Directory of Open Access Resources (DOAR) can serve as a valuable resource in identifying which journals are OA.[63]
2. *Use existing standards for identifying OA content* in a machine-readable form. The National Information Standards Organization (NISO) has recently closed a round of comments on a proposed standard for OA metadata and indicators.[64] The resulting standards should be adopted as widely as possible by all those publishing or otherwise distributing content. This applies to journals and repositories, but also to any content published, even non-institutional venues.
3. *Aggregate or otherwise provided federated search of institutional and subject repositories* at the national and regional levels. There are hundreds of repositories making OA content available from developing regions. Much of this content is not available as OA elsewhere (i.e., pre- and post-prints of articles published in toll journals), but current practices do not allow for tracking this output on a national or regional scale. Projects like LA Referencia in Latin America are making progress to

63 This service was launched after the initial stages of this chapter and so it was not included in the analysis. It can be found at: http://road.issn.org

64 http://www.niso.org/apps/group_public/download.php/12047/rp–22–201x_OA_indicators_draft_for_comments.pdf

this effect and its progress should be followed and, if successful, replicated elsewhere.

4. *Include non-traditional research outputs* in repositories so that they can be identified and tracked along with traditional research outputs.

7.3. REACH

Latin America is unique in its approach to OA and scholarly communications. The initiatives highlighted in this chapter represented a regional effort to improve access to scholarly works. This approach has served the region well in terms of improving the access and quality of scholarly communications in the region (Alperin, 2011). Importantly, the initiatives have resulted in a large amount of data being available to track scholarly production and OA in the region.

We recommend that other developing regions work on regional initiatives such as those of Latindex, SciELO, RedALyC, and CLACSO. In particular, we recommend the following:[65]

1. *Adopting regional, simple, clear and well-established editorial criteria* of what constitutes a high quality academic publication. Latindex, SciELO, and RedALyC in Latin America, and DOAJ internationally have already adopted this practice with varying degrees of success. Having regionally agreed-upon criteria ensures they are appropriate to the local context, consistent between countries, and provide journal editors with a standard they can strive towards to improve the quality of their publications.
2. *Create a directory and a catalog* of journals. A directory should be all-inclusive and the catalog should include those journals meeting the minimum specified in the regional editorial criteria. Directories and Catalogs such as Latindex's provide much needed context for all other indicators that may be produced.
3. *Work with a regional cooperative model* that is decentralized with national focal points and involvement from representatives from as many countries as possible. This is a relatively low-cost effort to coordinate and distributes the effort of col-

65 This is not to say that other aspects of the work of these initiatives are not to be praised or worthy of recommendation. This chapter is focused on OA indicators and therefore only comments on the activities of these initiatives as they pertain to the creation and usefulness of indicators of scholarly activity in general and OA in particular.

lecting information on publications to agencies that are closer to the place of publication.

4. *Aggregate content* from developing regions in regional portals in the manner of SciELO, and RedALyC. These initiatives provide visibility, improve indexing, and give infrastructure and support to make contents OA.
5. *Index and normalize metadata* in order to produce bibliographic and bibliometric indicators at journal, institution, and country level. SciELO and RedALyC both do this, with different foci, and benefit from the critical mass of journals and articles necessary for the indicators to be meaningful.
6. *Expose as many statistics as possible* including, growth, usage, and altmetrics. CLACSO currently shares its collection size over time and more recently started offering usage statistics at the article level. Ideally, these statistics would be downloadable and potentially exposed over an API or in ways compatible with federated repositories (like those proposed by LA Referencia). There are many challenges as advantages of metrics in repositories (Konkiel, 2013), but we recommend erring on the side of more metrics while the challenges are overcome.
7. *Begin exploration of both article-level metrics and altmetrics* in the way that SciELO and some journal portals from around the region are doing. The value and significance of these metrics in the long-term is still unknown, but their increased presence and use in developing regions will allow a greater understanding and facilitate the creation of new indicators of growth, impact, and reach.

7.4. IMPACT

Impact is an elusive concept, even when a limited definition is employed, such as the one proposed here. Traditionally, this term has been applied by the bibliometrics and innovation studies communities to mean "citation impact," a concept that, because of the limitations in counting citations, has always held developing regions at a disadvantage. In this chapter, we offer a broader conceptualization of impact; one that can take alternative forms, including some that cannot be counted through any metrics. We seek to present the existing indicators of impact, from traditional citation measures to alternative metrics, but remind the reader that impact should never be defined by what can currently be counted.

That said, two major developments are contributing to better indicators of impact of research from developing regions. The first is the

development of SciELO that, for the first time, has provided a more accurate count of citations of Latin American research (and more recently of South African research). As was detailed above, this method is not without its limitations, but it is a step in the right direction. The second has been the development of altmetrics, which has expanded the bibliometrics community's notion of impact by providing a quantifiable way of capturing *some form* of impact. It is early days in the field of altmetrics and we cannot be certain what form of impact is captured by these metrics. Yet, there is tremendous potential. To increase our ability to track impact of Latin American research (and potentially increase its impact), we recommend the following:

1. *Use the broadest definition of impact* that applies in a given context. Continue to explore alternative indicators of impact, but be careful to distinguish between the form of impact of interest and the metrics that can be collected with existing tools.
2. *Support the development of open altmetric tools* to improve the tracking of research from developing regions. It may be advantageous to support altmetric providers that cater to the developing region context.
3. *Research on altmetrics using content from developing regions* to give new insights onto the potential of *altmetrics*. There is a complete lack of data on altmetrics from developing regions, making it impossible to verify if existing claims about impact hold true in the developing context.
4. *Expand citation indexing* to journals outside of SciELO. While the work that SciELO has been doing is invaluable, and indexing citations is costly, the tracking of citations provides an invaluable measure of academic impact. As the existing altmetric studies also show, citations are also a valuable benchmark for other metrics because the properties of citations has been studied in detail over several decades
5. *Take caution when using any metric of impact* and do not confuse citation impact or other dimensions of impact from *altmetrics* are indicators of economic, policy, personal, and professional-practice impact. Taylor (2013) provides a useful starting point for understanding some of the difficulties in using *altmetrics* for capturing social impact.

8. CONCLUSIONS

After the careful consideration of all the data and resources reviewed to produce this chapter, two major trends emerge. First, the use, ap-

proach, and models of OA in developing regions generally, and Latin America in particular, have been extremely positive and should be strengthened and expanded. There are important lessons to be learned from documenting and assessing OA scholarly communication in developing regions, especially as they relate to generating robust OA indicators. Second, it is key to improve the quality of available OA indicators, with the hope that better indicators will help strengthen scholarly communications in developing regions, and thereby contributing in substantial way to removing the "developing" qualifier.

This is a lofty goal, but one that developing regions have embraced by supporting most of the research and development as well as scientific communications with public funds. Unlike the global North, developing regions have not ceded the responsibilities of communicating scholarship to for-profit commercial publishers. So, while in the United States and Europe debate between "green OA" and "gold OA," and whether to adopt author-pays models, regions like Latin America are expanding their research production improving the quality of the publications and starting to produce alternative indicators of access, reach, and quality of scientific communications through non-for profit publicly funded models. Despite the noticeable contributions and the local and regional improvements, developing regions have been noticeably absent from the global conversation around OA.

This chapter wants to call the attention to this absence, by contributing to the improvement of our understanding of OA in developing regions, and by voicing the importance of recognizing the contributions of developing regions to the global conversation around OA. We have sought to do this in two ways: first, by presenting a careful analysis of existing and desired OA indicators, we hope to equip scholars, policymakers, and others to better track and assess the growth, impact, and reach of scientific research from developing regions. An increased understanding of the impact and reach of research from developing regions would not only be useful within the regions, but also in determining and demonstrating which aspects of the models and approaches employed by developing regions are successful and worthy of replication. Second, by providing relevant information needed for improving the quality of the OA indicators themselves, we hope to decrease the unfavorable reliance of for-profit restricted commercial indexes that have been traditionally biased against developing regions. Investing in the generation of more robust and comprehensive OA indicators from developing regions has the potential to be a viable alternative to for-profit restricted commercial databases, which in turn will allow bibliometric studies to be inclusive of research, researchers, and institutions from the developing world.

As is highlighted in this chapter, developing regions using several strategies, and overcoming many difficulties have moved in a positive direction in providing appropriate tools for enhancing as well as understanding the growth, impact, and reach of their research. This chapter has described and analyzed a variety of such measures, both from regionally produced indicators and from indicators produced in the global North. As discussed in this chapter, a combination of both is necessary.

Developing regions, especially Latin America, have much to contribute to the global discussion around OA, not only OA indicators, but also in terms of scientific content, cooperative models, and best practices. By increasing and improving the OA indicators available from developing regions, we hope to have provided a valuable starting point for entering in these discussions.

9. REFERENCES

Adams, C. (2013, December). Open access in Latin America: Embraced as key to visibility of research outputs. *SPARC news*. Retrieved from http://www.sparc.arl.org/news/open-access-latin-america-embraced-key-visibility-research-outputs

Alonso-Gamboa, J. O., & Russell, J. M. (2012). Latin American scholarly journal databases: a look back to the way forward. *Aslib Proceedings*, *64*(1), 32–45. doi:10.1108/00012531211196693

Alperin, J. P. (2011). Academic publishing in a global context: The politics and practices of publishing in English (review). *Journal of Scholarly Publishing*. Retrieved from http://scholar.google.com/scholar?hl=en\&q=juan+pablo+alperin\&btnG=Search\&as_sdt=0,5\&as_ylo=\&as_vis=0\#5

Alperin, J. P. (2013a). Ask not what altmetrics can do for you, but what altmetrics can do for developing countries. *Bulletin of the American Society for Information Science and Technology*, *39*(4), 18–21. doi:10.1002/bult.2013.1720390407

Alperin, J. P. (2013b). What it means for PKP to offer article level metrics. In *Scielo 15*. Sao Paolo. Retrieved from https://speakerdeck.com/jalperin/what-it-means-for-pkp-to-offer-article-level-metrics

Alperin, J. P. (2013c, October). Are ALMs/altmetrics propagating global inequality? *Article level metrics workshop 2013*. Retrieved from https://speakerdeck.com/jalperin/altmetrics-propagating-global-inequality

Alperin, J. P., Fischman, G. E., & Willinsky, J. (2011). Scholarly communication strategies in Latin America's research-intensive

universities. *Revista Educación Superior Y Sociedad*, *2*(16). Retrieved from http://ess.iesalc.unesco.org.ve/index.php/ess/article/viewArticle/409

Alperin, J. P. (2014). South America: Citation databases omit local journals. *Nature*, 511(7508), 155–155. doi:10.1038/511155c

Antelman, K. (2004). Do open-access articles have a greater research impact. *College & Research Libraries*, *65*, 372–382.

Babini, D. (2012). Scientific output from Latin America and the Caribbean – Identification of the Main Institutions for Regional Open Access Integration Strategies. *SSRN Electronic Journal*. doi:10.2139/ssrn.2125996

Babini, D. (2013). Open access in Latin America and the Caribbean (LAC). In *Report of the regional Latin American and Caribbean consultation on open access to scientific information and research – concepts and policies*. Kingston. Retrieved from http://es.slideshare.net/Babini/clacso-dominique-babini-open-access-in-latin-america

Bakkalbasi, N., Bauer, K., Glover, J., & Wang, L. (2006). Three options for citation tracking: Google Scholar, Scopus and Web of Science. *Biomedical Digital Libraries*, *3*, 7. doi:10.1186/1742-5581-3-7

Bar-Ilan, J. (2012). JASIST 2001-2010. *Bulletin of the American Society for Information …*, 24–28. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/bult.2012.1720380607/full

Bar-Ilan, J., Haustein, S., Peters, I., Priem, J., Shema, H., & Terliesner, J. (2012). Beyond citations: Scholars' visibility on the social Web. In *Proceedings of the 17th international conference on science and technology indicators*. Montreal. Retrieved from http://arxiv.org/abs/1205.5611

Björk, B.-C. (2011). A study of innovative features in scholarly open access journals. *Journal of Medical Internet Research*, *13*(4), e115. doi:10.2196/jmir.1802

Björk, B.-C., Welling, P., Laakso, M., Majlender, P., Hedlund, T., & Gudnason, G. (2010). Open access to the scientific journal literature: situation 2009. *PloS One*, *5*(6), e11273. Retrieved from http://dx.plos.org/10.1371/journal.pone.0011273

Bollen, J., & Sompel, H. V. de. (2008). Usage impact factor: The effects of sample characteristics on usage-based impact metrics. *Journal of the American Society for Information Science and Technology*, *59*(1), 136–149. doi:10.1002/asi.20746

Bollen, J., Rodriguez, M., & Sompel, H. V. de. (2007). MESUR: usage-based metrics of scholarly impact. In *Proceedings of the*

7th aCM/iEEE-cS joint conference on digital libraries. Retrieved from http://dl.acm.org/citation.cfm?id=1255273

Bollen, J., Van de Sompel, H., Hagberg, A., & Chute, R. (2009). A principal component analysis of 39 scientific impact measures. *PloS One*, *4*(6), e6022. doi:10.1371/journal.pone.0006022

Bollen, J., Van de Sompel, H., Rodriguez, M. A., Sompel, H. V. D., & Alamos, L. (2008). Towards usage-based impact metrics. In *Proceedings of the 8th aCM/iEEE-cS joint conference on digital libraries - jCDL '08* (p. 231). New York, New York, USA: ACM Press. doi:10.1145/1378889.1378928

Brody, T., & Harnad, S. (2004). Comparing the impact of open access (OA) vs. non-OA articles in the same journals. D-Lib Magazine, 10 (6). *ECS EPrints Repository*. Journal article. Retrieved from http://eprints.ecs.soton.ac.uk/10207/1/06harnad.html

Carvalho Neto, S. (2013, May). Scientific Serials in the Capes Qualis and the Brazilian Open Access Journals. *PKP scholarly publishing conference 2013*. Retrieved from http://pkp.sfu.ca/pkp2013/paper/view/366/267

Cetto, A. M., & Alonso-Gamboa, J. O. (1998). Scientific periodicals in Latin America and the Caribbean: a global perspective. *Interciencia*, *23*(525), 84–93. Retrieved from http://www.interciencia.org/v23_02/cetto.pdf

Charles W. Bailey, J. (2013). Altmetrics Bibliography. Retrieved from http://digital-scholarship.org/alt/altmetrics.htm

Collazo-Reyes, F. (2013). Growth of the number of indexed journals of Latin America and the Caribbean: the effect on the impact of each country. *Scientometrics*, *98*(1), 197–209. doi:10.1007/s11192-013-1036-2

Collazo-Reyes, F., Luna-Morales, M. E., Russell, J. M., & Pérez-Angón, M. (2008). Publication and citation patterns of Latin American & Caribbean journals in the SCI and SSCI from 1995 to 2004. *Scientometrics*, *75*(1), 145–161. doi:10.1007/s11192-007-1841-6

Costas, R., Zahedi, Z., & Wouters, P. (2014). Do altmetrics correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. Retrieved from http://arxiv.org/abs/1401.4321

COUNTER. (2012). The COUNTER code of practice for e-resources. *Counter Online Metrics* (No. April 2012) (pp. 1–29). Retrieved from http://www.projectcounter.org/r4/COPR4.pdf

Davis, P. M., & Price, J. S. (2006). eJournal interface can influence usage statistics: Implications for libraries, publishers, and

Project COUNTER. *Journal of the American Society for Information Science and Technology*, *57*(9), 1243–1248. doi:10.1002/asi.20405

De Volder, C., Babini, D., López, F. A., Aguado López, E., Becerril García, A., Rogel Salazar, R., & Arzuaga Magnoni, J. (2013, January). Informe sobre la producción científica de Argentina en revistas iberoamericanas de acceso abierto en redalyc.org 2005-2011. Retrieved from http://ri.uaemex.mx/xmlui/handle/123456789/573

Edgar, B. D., & Willinsky, J. (2010). A survey of the scholarly journals using open journal systems. *Scholarly and Research Communication*, *1*(2), 1–22.

Eysenbach, G. (2006). Citation advantage of open access articles. *PLoS Biology*, *4*(5), e157. doi:10.1371/journal.pbio.0040157

Eysenbach, G. (2011). Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact. *Journal of Medical Internet Research*, *13*(4), e123. doi:10.2196/jmir.2012

Fausto, S., Machado, F. A., Bento, L. F. J., Iamarino, A., Nahas, T. R., & Munger, D. S. (2012). Research blogging: indexing and registering the change in science 2.0. *PloS One*, *7*(12), e50109. doi:10.1371/journal.pone.0050109

Gantman, E. R. (2011). La productividad científica argentina en Ciencias Sociales: Economía, Psicología, Sociología y Ciencia Política en el CONICET (2004-2008). *Revista Española de Documentación Científica*, *34*(3), 408–425. doi:10.3989/redc.2011.3.829

Garfield, E. (1983a). Mapping science in the third world. *Science and Public Policy*, (June), 112–127. Retrieved from http://spp.oxfordjournals.org/content/10/3/112.full.pdf+html

Garfield, E. (1983b). Science in the third world. Whether or not a third world article is highly-cited depends on the recognition it gets from authors in the developed countries. *Science Age*, *1*(3&4), 59–65. Retrieved from http://garfield.library.upenn.edu/papers/scienceagep59y1983.pdf

Garfield, E. (1996). Significant scientific literature appears in a small core of journals. *The Scientist*, *10*(17). Retrieved from http://cat.inist.fr/?aModele=afficheN\&cpsidt=10620978 http://www.garfield.library.upenn.edu/commentaries/tsv10(17)p13y090296.html.

Garfield, E., & Sher, I. H. (1963). New factors in the evaluation of scientific literature through citation indexing. *American Documentation*, *14*(3), 195–201. doi:10.1002/asi.5090140304.

Gasca Pliego, E., Martínez Carbajal, R., Tinoco García, I., Arriaga Jordán, C., González López, S., & Aguado López, E. (2013). Informe sobre la producción científica de la UAEMEX en revistas iberoamericanas de acceso abierto en redalyc.org, 2005-2011. Universidad Autónoma del Estado de México. Retrieved from http://ri.uaemex.mx/xmlui/handle/123456789/205

GOAP. (2013). Global open access portal. United Nations Educational, Scientific and Cultural Organization (UNESCO). Retrieved from http://www.unesco.org/new/en/communication-and-information/portals-and-platforms/goap

Gómez, N., Bustos-Gonzalez, A., Santillan-Aldana, J., & Arias, O. (2009). Open access indicators and information society: the Latin American case. *OCLC Systems & Services*, *25*(2), 82–92. doi:10.1108/10650750910961884

Guédon, J.-C. (2008). Mixing and matching the green and gold roads to open access—Take 2. *Serials Review*, *34*(1), 41–51. doi:10.1016/j.serrev.2007.12.008

Haider, J. (2005). *The geographic distribution of open access journals*. Salvador/Bahia, Brazil. Retrieved from http://www.soi.city.ac.uk/~ch696/icml9_poster_haider.pdf

Harnad, S., Brody, T., Vallières, F., Carr, L., Hitchcock, S., Gingras, Y., … Stamerjohanns, H. (2004). The access/impact problem and the green and gold roads to open access. *Serials Review*, *30*(1), 310–314. Retrieved from http://www.sciencedirect.com/science/article/pii/S0098791307001530

Haustein, S., & Peters, I. (2013). Tweeting biomedicine: an analysis of tweets and citations in the biomedical literature. *Journal of the Association for Information Science and Technology*, 1–22. doi:10.1002/asi.23101

Herb, U. (2010). Open access statistics: Alternative impact measures for open access documents? An examination how to generate interoperable usage information from distributed open access services. *L'information Scientifique Et Technique Dans L'univers Numérique. Mesures Et Usages*, 165–178. Retrieved from http://eprints.rclis.org/19068/1/preprint.pdf

Hitchcock, S. (2013, June). The effect of open access and downloads ('hits') on citation impact: a bibliography of studies. University of Southampton. Retrieved from http://eprints.soton.ac.uk/354006/1/oacitation-biblio-snapshot0613.html

Irvine, J., & Martin, B. R. (1984). *Foresight in science: Picking the winners* (p. 166). Pinter Pub Ltd. Retrieved from http://

www.amazon.com/Foresight-Science-Picking-John-Irvine/dp/086187496X

Jacsó, P. (2012). Using Google Scholar for journal impact factors and the h-index in nationwide publishing assessments in academia – siren songs and air-raid sirens. *Online Information Review*, *36*(3), 462–478. doi:10.1108/14684521211241503

Konkiel, S. (2013). Altmetrics in institutional repositories. Retrieved from https://scholarworks.iu.edu/dspace/handle/2022/17122

Laakso, M., & Björk, B.-C. (2012). Anatomy of open access publishing: a study of longitudinal development and internal structure. *BMC Medicine*, *10*, 124. doi:10.1186/1741-7015-10-124

Laakso, M., Welling, P., Bukvova, H., Nyman, L., Björk, B.-C., & Hedlund, T. (2011). The development of open access journal publishing from 1993 to 2009. *PloS One*, *6*(6), e20961. doi:10.1371/journal.pone.0020961

Lawrence, S. (2001). Free online availability substantially increases a paper's impact. *Nature*. Retrieved from http://www.nature.com/nature/journal/v411/n6837/full/411521a0.html

Leetaru, K., Wang, S., Cao, G., Padmanabhan, A., & Shook, E. (2013, April). Mapping the global Twitter heartbeat: The geography of Twitter. *First Monday*. Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/4366/3654

Leydesdorff, L., & Cozzens, S. (1993). The delineation of specialties in terms of journals using the dynamic journal set of the SCI. *Scientometrics*, *26*(1), 135–156. Retrieved from http://www.akademiai.com/index/RT705N2280353555.pdf

Leydesdorff, L., & Meyer, M. (2006). Triple helix indicators of knowledge-based innovation systems. *Research Policy*, *35*(10), 1441–1449. doi:10.1016/j.respol.2006.09.016

Leydesdorff, L., Cozzens, S., & Besselaar, P. V. D. (1994). Tracking areas of strategic importance using scientometric journal mappings. *Research Policy*, *23*(2), 217–229. Retrieved from http://www.sciencedirect.com/science/article/pii/004873339490054X

Li, X., & Thelwall, M. (2012). F1000, Mendeley and traditional bibliometric indicators. In *Proceedings of the 17th international conference on science and technology indicators*. Montrèal. Retrieved from http://2012.sticonference.org/Proceedings/vol2/Li_F1000_541.pdf

Li, X., Thelwall, M., & Giustini, D. (2012). Validating online reference managers for scholarly impact measurement. *Scientometrics*. Retrieved from http://link.springer.com/article/10.1007/s11192-011-0580-x

Lin, J., & Fenner, M. (2013). Altmetrics in evolution: defining & redefining the ontology of article-level metrics. *Information Standards Quarterly*, *25*(2), 20–26. Retrieved from http://www.niso.org/publications/isq/2013/v25no2/lin/

Liu, J., & Adie, E. (2013, May). Altmetric: Getting started with article-level metrics. figshare. doi:10.6084/m9.figshare.709018

Luna-Morales, M. E., & Collazo-Reyes, F. (2007). Historic and bibliometric analysis of the Latin American and Caribbean journals in the international Science Indexes: 1961-2005. *Revista Española de Documentación Científica*, *30*(4), 523–543. doi:10.3989/redc.2007.v30.i4.403

Lundvall, B. (1988). Innovation as an interactive process: from user-producer interaction to the national system of innovation. *Technical Change and Economic Theory*. Retrieved from http://www.ajstid.com/abstractlundvall.pdf

MacMillan, D., & Wakabayashi, D. (2013, December). Apple buys Twitter analytics startup Topsy Labs. *Wall Street Journal*. Retrieved from http://online.wsj.com/news/articles/SB10001424052702304854804579234450633315742

Marin, A., & Petralia, S. (2013). *Project proposal: Evaluating the impact of Latin American science within academia and beyond* (Vol. 1). Retrieved from http://flacso.org.br/oa/wp-content/uploads/2013/11/Convocatoria-Concurso-Projetos-2013-FLACSObr-PROJECT-CENIT.pdf

Mayr, P. (2006). Constructing experimental indicators for open access documents. *Research Evaluation*, *15*(2), 127–132. doi:10.3152/147154406781775940

McVeigh, M. (2004). Open access journals in the ISI citation databases: Analysis of impact factors and citation patterns: A citation study from Thomson Scientific. *Thomson Corporation*, *17*(October). Retrieved from http://biblioteca.uned.es/lenya/bibliuned/search-authoring/docpdf/oacitations2.pdf

Meho, L. I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of science versus scopus and google scholar. *Journal of the American Society for Information Science and Technology*, *58*(13), 2105–2125. doi:10.1002/asi.20677

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*, 529–538. Retrieved from http://dx.doi.org/10.1007/s11192-006-0168-z

Michels, C., & Schmoch, U. (2012). The growth of science and database coverage. *Scientometrics*, (October 2011). doi:10.1007/s11192-012-0732-7

Miguel, S., Chinchilla-Rodriguez, Z., & Moya-Anegón, F. de. (2011). Open access and Scopus: A new approach to scientific visibility from the standpoint of access. *Journal of the American Society for Information Science and Technology*, *62*(April), 1130–1145 ST–Open access and Scopus: A new appr. doi:10.1002/asi.21532

Moed, H. F. (2009). New developments in the use of citation analysis in research evaluation. *Archivum Immunologiae Et Therapiae Experimentalis*, *57*(1), 13–8. doi:10.1007/s00005-009-0001-5

Molloy, M. (2005). The internet in Latin America: development and reference sources. *Journal of Library Administration*, 129–148. doi:10.1300/J111v43n03

Nelson, R. R. (1993). National Innovation Systems: A Comparative Analysis. Retrieved from http://papers.ssrn.com/abstract=1496195

Nelson, R. R., Policy, N. Y. U. G. S. of B. A. C. for S., & Technology. (1982). *Government and technical progress: a cross-industry analysis* (p. 498). Pergamon Press. Retrieved from http://books.google.com.mx/books/about/Government_and_technical_progress.html?id=8V6sAAAAIAAJ\&pgis=1

Packer, A. L., & Meneghini, R. (2007). Learning to communicate science in developing countries. *Interciencia*, *32*, 643–647. Retrieved from http://www.scielo.org.ve/scielo.php?script=sci_arttext\&pid=S0378-18442007000900014\&lng=en\&nrm=iso\&ignore=.html

Piwowar, H. (2013). Introduction altmetrics: What, why and where? *Bulletin of the American Society for Information Science and Technology*, *39*(4), 8–9. doi:10.1002/bult.2013.1720390404

Priem, J. (2010, September). I like the term #articlelevelmetrics, but it fails to imply *diversity* of measures. Lately, I'm liking #altmetrics. [Tweet]. Retrieved from https://twitter.com/jasonpriem/status/25844968813

Priem, J., Piwowar, H., & Hemminger, B. M. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. Retrieved from http://jasonpriem.org/self-archived/PLoS-altmetrics-sigmetrics11-abstract.pdf

Rosenzweig, P., Rosales López, N. F., Rodríguez Contreras, F., Bauste Sarache, M. T., Guzmán Useche, E., Tinoco García, I., … Gutiérrez Cortés, P. M. (2013, March). Informe sobre la producción científica de Venezuela en revistas iberoamericanas

de acceso abierto en redalyc.org, 2005-2011. Universidad de Los Andes. Retrieved from http://ri.uaemex.mx/xmlui/handle/123456789/708

Shema, H., Bar-Ilan, J., & Thelwall, M. (2012). Research blogs and the discussion of scholarly information. *PloS One*, *7*(5), e35869. doi:10.1371/journal.pone.0035869

Shepherd, P. T. (2011). PIRUS2: individual article usage statistics – developing a practical, global standard. *Learned Publishing*, *24*(4), 279–286. doi:10.1087/20110405

Shuai, X., Pepe, A., & Bollen, J. (2012). How the scientific community reacts to newly submitted preprints: article downloads, Twitter mentions, and citations. *PloS One*, *7*(11), e47523. doi:10.1371/journal.pone.0047523

SPARC. (2013). Article-Level Metrics. *SPARC Primer*. Retrieved from http://www.sparc.arl.org/initiatives/article-level-metrics

Taylor, M. (2013, June). The challenges of measuring social impact using altmetrics. *Research Trends*. Retrieved from http://www.researchtrends.com/issue-33-june-2013/the-challenges-of-measuring-social-impact-using-altmetrics/?

Testa, J. (2011). The globalization of Web of Science: 2005-2010. Retrieved from http://thomsonreuters.com/products_services/science/free/essays/regional_content_expansion_wos/

Testa, J. (2012). The Thomson Reuters journal selection process. Retrieved from http://thomsonreuters.com/products_services/science/free/essays/journal_selection_process/

The San Francisco Declaration on Research Assessment (DORA). (2012). Retrieved from http://am.ascb.org/dora/

Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, C. R. (2013). Do altmetrics work? Twitter and ten other social web services. *PloS One*, *8*(5), e64841. doi:10.1371/journal.pone.0064841

Torres-Salinas, D., Cabezas-Clavijo, Á., & Jiménez-Contreras, E. (2013). Altmetrics: New indicators for scientific communication in Web 2.0. *Comunicar*, *21*(41), 53–60. doi:10.3916/C41-2013-05

UNESCO. (2013). Report of the regional Latin American and Caribbean consultation on Open Access to scientific information and research – concepts and policies. Retrieved from http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/news/report_open_access_en.pdf

Velho, L. (2004). *Science and technology in Latin America and the Caribbean: an overview* (No. February). United Nations University Institute for New Technologies. Retrieved from http://ideas.repec.org/p/dgr/unuint/200404.html.

Vessuri, H., Guedon, J.-C., & Cetto, A. M. (2013). Excellence or quality? Impact of the current competition regime on science and scientific publishing in Latin America and its implications for development. *Current Sociology*. doi:10.1177/0011392113512839

Villanueva, D. (2013). Comparativo de Indicadores Bibliométricos Latinoamericanos: Los casos de BIBLAT, RedALyC, SciELO. Retrieved from http://bidigital.blogspot.mx/2013/09/comparativo-de-indicadores.html

Wouters, P., & Costas, R. (2012). *Users, narcissism and control: Tracking the impact of scholarly publications in the 21st century*. Retrieved from http://sticonference.org/Proceedings/vol2/Wouters_Users_847.pdf

SECTION II

THE METRICS USED IN THE DEVELOPMENT OF THE SCIELO NETWORK MODEL

Abel L. Packer*

1. INTRODUCTION

The SciELO Network is the result of an international cooperative program on scientific communication established in 1998 to advance research published by independent, national, open-access journals. In 2014, the network covers approximately 1,000 active journals gathered in national collections from 16 countries from Latin America, Portugal, Spain, and South Africa (SciELO, 2014). The network publishes about 40,000 new articles per year and has accumulated over 450,000 over its 16 years of operation. The SciELO Program follows the Open Access Golden model.

This chapter presents an overview of the SciELO Program objectives and strategy and the challenge it faces in adopting and using the classical, as well as the new web-based, metrics and their adequacy to tackle the performance of SciELO Network collections and the individual journals it indexes and publishes.

2. OVERVIEW OF THE SCIELO PROGRAM

The principles, objectives, and strategies of the SciELO Program are implemented through the indexing, publishing, and interoper-

* Director of SciELO / FAPESP Program.

ating of peer-reviewed journals that are gathered in national collections according to selection criteria defined and applied by nationally established scientific committees and operated by a common methodological and technological platform. These collections are managed and funded at the national level. This decentralized approach contributes to developing national capacities and infrastructures in scholarly communication. The specific objectives are to help increase the visibility, availability, quality, use, impact, and credibility of individual journals and the research they communicate. One of the main reasons for the creation of SciELO was to complement the international indexes on both coverage of journals and measurement of their impact (Packer, 2009). Particular importance is given to locally published journals that are not covered by international indexes, as well as offering research agencies, institutions, journals, and authors additional sources of indexing and evaluation of journals.

To accomplish its objectives, the SciELO Program proposes to develop each national collection according to the following general characteristics:

- There is one collection per country.
- It is developed and operated as an integral part of national research system;
- It is led and funded by national research agencies.
- It is managed directly by national research agencies or by leading national research institutions.
- Collections encompass three main functions that are offered online:
  - The first function is Journal Indexing that includes the establishment of bibliographic control of the relevant research communicated via journals published by national institutions, and the evaluation of journals' performance according their presence, influence, and impact measured by number of citations, references, and downloads received by the articles.
  - The second function is Journal Publishing that includes the online availability and unrestricted open access of publications through a portal that provides functionalities for searching and navigation through journals, articles, and metrics of usage and impact.

  - The third function is Journal Interoperability that includes the availability and interchange of journal and articles' metadata with bibliographic indexes and systems.
- The journal collections are developed under the assistance of a scientific committee representative of the national research community and responsible for the application of the journal evaluation and indexing criteria. The committee is responsible for the adaptation of the SciELO Network's common evaluation criteria to the country research conditions.
- All collections follow the same principles, methodologies, and technologies. Collections are created following a pilot of 3 to 5 journals that evolve to a public "in development" collection integrated to the network until reaching the "certified" status as compliant with the SciELO principles, methodologies, and technologies. The pace of development and the level of updating vary significantly among the collections.
- The collections' national governance and operation are decentralized but integral part of the SciELO Network.
- Collections contribute to the development of national research policies particularly regarding funding and evaluation of journals.
- The main principles of the SciELO Network operation are the following: (a) development with focus on improvement of the quality of contents and processes; (b) adherence to the Open Access Golden model with attribution CC-BY-NC or CC-BY or equivalent; (c) coverage of all disciplines; (d) multilingualism; (e) independence of journal editorial policies and management; and, (f) online interoperability.

At the network level, the SciELO program has been led by the SciELO Brazil collection that operates under the São Paulo Research Foundation (FAPESP) as a special research program, and is also supported by the Brazilian National Council for Scientific and Technological Development (CNPq). SciELO Brazil is also responsible for the maintenance of the methodological and technological platform, and operates as the Secretariat of the Network, which includes the organization of periodical meetings, the management of the online collaborative space, and ensuring compliance to the SciELO principles, methodologies, and technologies. The SciELO methodological and technological platform follows the Virtual Health Library (VHL) scientific information model, developed by the Latin American and Caribbean Center

on Health Sciences Information, known by its acronym BIREME of its original name Regional Library of Medicine in Portuguese. BIREME is a center of the Pan American Health Organization World Health Organization (PAHO/WHO), which focuses on the democratization of scientific information through network-based information and knowledge management (BIREME, 2011).

SciELO Network representatives gathered on October 22, 2013, in São Paulo as part of the SciELO 15 Years Celebration Conference and agreed to renew and advance the program through the following main decisions (SciELO, 2013):

- Three main lines of actions will drive the program the next three years: professionalization, internationalization, and sustainability of journal management and operation. These lines of action will be implemented according national conditions and priorities;
- SciELO open access platform will be improved through the provision of common services covering all the steps of journal publishing workflow;
- SciELO national collections will operate or act as intermediary for editing and publishing services envisage to follow state of the art policies and procedures; and,
- SciELO Network, SciELO Collections, and SciELO journal performance will be evaluated in a systematic way using a set of metrics covering number of citations, online presence, and number of downloads.

Nationally published journals that comply with the conditions SciELO demands for indexing and publishing are required to present, or present in the near future, the following general characteristics:

- Well-established editorial policies, governance, and manuscript peer-reviewed processing flow in compliance with editorial and publishing standards;
- Well-defined mission and thematic scope considering that most of the journal have a general coverage of their disciplines and topics;
- Language of publication is a journal policy and decision of their editors. SciELO has a multilingual approach with emphasis on the usage of English to improve internationalization. Journals publishing articles in different languages should

include the option of simultaneously publishing in more than one language;

- Journals should adhere to SciELO rules and recommendations; and,
- Journals are expected to progressively increase their performance to remain in SciELO collections.

However, the development of SciELO journals face common and well-known challenges that manifest with different nuances among the countries: lack of well-established national and institutional policies that effect the stable financing of the journals, lack of professionalism on the production of the journals, and the lack of internationalization. The SciELO Program has made overcoming these barriers a top priority over the next three years. Again, journal-related metrics play a crucial role to understanding journal performance over time and the SciELO Program as whole, as well as the implementation of SciELO's new lines of action towards professionalization and internationalization.

3. SCIELO MODEL: HOW IT OPERATES

The SciELO Network expresses itself through a common global portal, www.scielo.org, which provides access to the whole content of the network, the individual collections, and the usage and impact metrics. The portal and the individual collections are updated weekly. Each collection is responsible for keeping its database current. A weekly harvesting system collects new contents from each collection and updates the global portal database. Also weekly, SciELO shares and/or transfers new deposits to the main indexes, including Web of Science (WoS), Scopus, Google Scholar, CROSSREF, PubMed, LILACS, and AGRIS, among others. Dozens of web-based indexing robots also harvest the SciELO database contents daily.

National collections are updated according the following procedures:

- SciELO indexed journals are managed and operated independently, and agree to send the contents to the SciELO national coordinating unit to be aggregated in a national portal and published in open access without embargo;
- Journals are responsible for transferring their new contents to a national SciELO server;
- Journal contents are sent in different formats: PDF, WORD, HTML, or XML according SciELO DTD;

- All texts are marked up according a SciELO DTD. The marking up process is carried out either under the responsibility of the journal or the national coordinating center, using in-house resources or hiring private services; and,
- National databases comprise three main components: (a) the indexing database, including article metadata; (b) the full-text database of the marked-up texts and a repository of correspondent PDF files stored in the operating system file system; and, (c) the bibliometric database. They are updated weekly.

Since its pilot development in 1997, SciELO adopted the markup of the original texts as the source of all document metadata in order to guarantee consistency (Packer et al, 1998). Therefore, all metadata, bibliographic records, summaries, registry of collections, etc. are always produced at the time of text deposit or extracted from the original contents. SciELO original DTD was adopted from the ISO 8879-1986 (SGML - Standard Generalized Markup Language) and ISO 12083-1994 (Electronic Manuscript Preparation and Markup). When SciELO was initially created XML and the related formatting and displaying tools did not exist. The corresponding markup process departs from the original text converted to HTML format and is carried out using Microsoft Word with an add-on that drives the markup process according to the DTD elements structure. The HTML format, according SciELO specification, is used to anticipate the final display of the texts. The original DTD was updated several times in the last 15 years and its current version is documented in the "About" section of the SciELO global portal. This DTD is oriented to the detailed markup of articles' front and back sections – keeping the body content in HTML format with a simple markup of beginning and end of paragraphs.

Since 2013, SciELO adopted the SciELO Publishing Schema derived from the standard ANSI/NISO Z39.96-2012: Journal Article Tag Suite (JATS) to extend the detailed markup to the article body. This new markup methodology will be implemented throughout the SciELO Network in the next several years, taking into consideration the managerial, financial, and technological complexities involved in the change. The objective is to produce the XML text as the reference source-digital-text of articles for preservation and generation of the other formats (HTML, PDF, ePUB, etc.). Under the new publishing schema, SciELO will also collect high definition images of the article figures. The adoption of the JATS DTD will bring several advancements: (a) better quality of textual data elements that will contribute to improved database operation, indexing, metrics calculation, and

interoperability; (b) full capability to display the texts in different formats and to adapt to the different fixed and mobile devices; and, (c) health sciences journals will be able to be indexed and stored in the PUBMED Central Repository.

4. SCIELO METHODOLOGIES TO GENERATE BIBLIOMETRIC INDICATORS

The generation of metrics associated to SciELO collections and journals is a functionality that is part of the original conception of SciELO and reflects SciELO's commitment to increasing the visibility, usage, and impact of its journals and the research they communicate. Thus, the creation of journal and article metrics is an integral part of the SciELO Methodology, following the three main functions of the program described before: indexing, online publishing, and interoperability. The indexing function is implemented through bibliographic and bibliometric databases to meet the following goals: (a) the retrieval of records using Boolean expression applied to inverted indexes of bibliographic elements such as words of title and abstract, authors, keywords, and journal title; (b) the production of bibliometric indicators; and, (c) the interoperability with other Web indexes and systems.

The bibliometric database records all bibliographic references cited by journal articles. By June 2014, the number of citations records is over 10.8 million. The bibliographic references records are generated from the markup process that identifies all data elements. This is the most complex and time-consuming component of the SciELO Methodology. As the records are extracted from the marked-up texts, they often need additional processing to further clean up the data. These processes include the checking of the correct structure of the bibliographic record, which varies according the standard adopted by the journals, and the identification of the journal titles cited, which comprises about 70% of the citations followed by books and other types of literature. The automated process that operates on predefined rules is able to successfully identify about 85% of the cited journals (Mugnaini, Tardelli & Tuesta, 2004) The improvement in the data cleaning process depends on the development of more sophisticated algorithms. The references database represents one of the most important information sources for SciELO Network countries and their contribution to national and international research.

SciELO Methodology also provides procedures to record article access via the SciELO interface. However, the application of this feature requires high availability of disk space to log all the transac-

tions. Due to the limitations of the hardware infrastructure of the national collections, these data logs are currently available in the SciELO Portal only for Brazil and Chile. The development of a more powerful system at the network-level to keep track and audit the access and downloads of contents throughout the decentralized collections is pending. This system is of crucial importance due to the fact that SciELO content is intensively used and evidence is needed to document the importance of nationally published journals in terms of influence and impact. A comprehensive, clean, and audited set of indicators on downloads would help balance the generally low performance of SciELO journals using Impact Factor on the international indexes.

From the bibliometric database, both the SciELO global portal and each collection interface generate and publish bibliometric indicators in two main ways: (a) online interactive interfaces updated weekly; and, (b) an exhaustive list of bibliometric indicators updated every six months. All these indicators are available online. In addition, the SciELO Program relies on the metrics provided by international indexes to follow the performance of the collections, journals, and articles.

4.1. SCIELO JOURNALS BIBLIOMETRIC INDICATORS INTERACTIVE INTERFACE

An online interactive interface used to extract journal-related bibliometric indicators is available at the collection- and network-levels. The interface allows the calculation of indicators for one or more journals.

The online interactive indicators are similar to those produced by Thomson Reuters Journal Citation Reports and they are accessible through the menu shown in Figure 1. It is a powerful mechanism to follow up the performance of individual journals within a SciELO collection. However, it is important to emphasize that the indicators' values correspond to the universe of SciELO journals only and therefore do not consider citations the journals received from articles published by journals not indexed by SciELO. To address this limitation, SciELO entered into an agreement with Thomson Reuters to operate the SciELO Citation Index (SciELO CI) within the Web of Science (WoS) platform as of January 2014. The first annual report of the citations granted to SciELO journals from both SciELO and WoS journals will be available in September 2014.

The indicators interface menu is accessible from the SciELO global portal (through the "SciELO in numbers" box) or from each collection portal (through the "Reports" submenu).

Figure 1
Interactive menu of journal citation indicators with links to SciELO Network

Library collection

Last processed date
Oct-28-2013

Journal citation reports

- Source data
- Impact factor on a two-year basis
- Impact factor on a three-year basis
- Half-life
- Received citations
- Granted citations

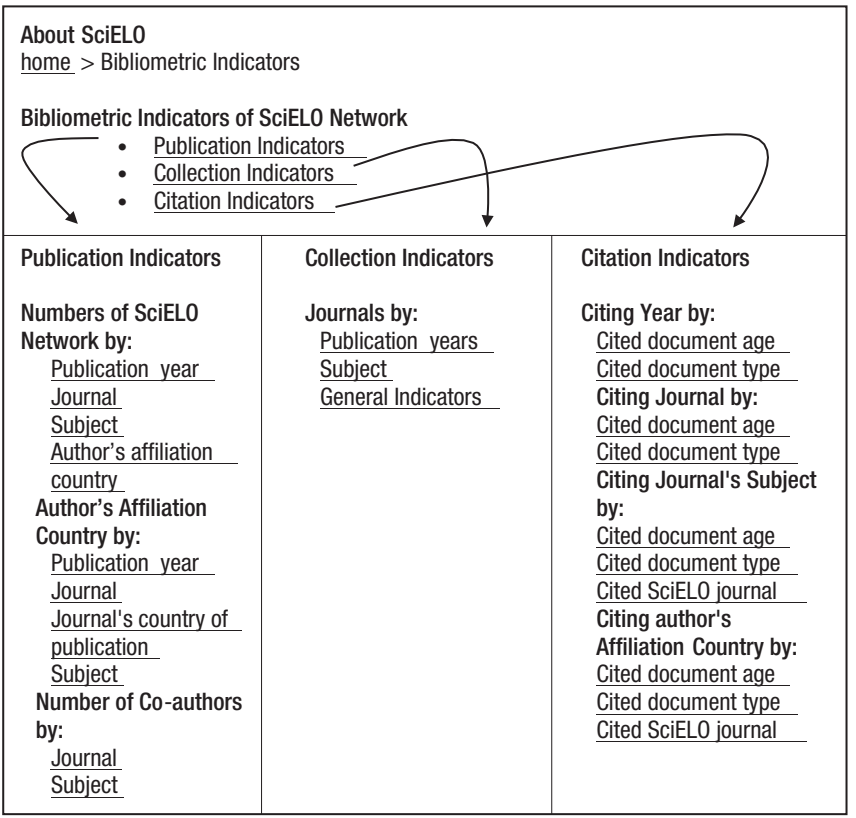
As the menu shown in Figure 1 illustrates, SciELO provides to journals editors, researchers, and general users a detailed and updated list of indicators on its journals' performance. Therefore, SciELO not only contributes to open access by communicating the research produced in nationally published journals, it also provides updated indicators that measure on-time publishing of journals, the number of articles published, the citations they received from other SciELO Journals, and citations they grant. These indicators contribute to the development of the collection as it accepts news journals into the collection or keeps those already indexed. The list of citations granted permits the identification of journals not yet in SciELO that receive a high number of citations and therefore are good candidates for indexing, as well as the universe of journals cited and the degree of national and international orientation of the research published.

4.2. SCIELO YEARLY-BASED BIBLIOMETRIC INDICATORS

SciELO publishes a list of over 20 annual-based bibliometric indicators that are updated every six months. The indicators are organized in Microsoft Excel tables ready to be downloaded by users and managed on local computers. These tables cover all SciELO journals from the calendar year 2000 onwards. The list is accessible from the "About" section of the SciELO Global portal. These Indicators are organized into three main groups:

- Publication Indicators – number of journals and articles from all the network;
- Collection indicators – number of articles and citations at journal individual and collection levels; and,
- Citation indicators – covers different distribution of the citations received by journals from all other network journals.
- The Figure 2 shows the menus sequence to access the SciELO annually based bibliometric indicators.

Figure 2
SciELO Network. Menus to access and download yearly-based bibliometric indicators



4.3. SCIELO INTERACTIVE INTERFACE FOR METRICS OF DOWNLOADS

The SciELO platform keeps track of all access to the journals and their articles, which allows for the measurement of the performance of the collections, journals, issues, and articles. The interface provides access by year and months so editors, authors, researchers, users, and the SciELO coordinators can continuously follow the performance of the journals and articles.

The logging of the access is carried out in each node of the navigation and requires high availability of disk space, which limits the online availability of the statistics. Because of this, many collections have limited the publication of the access statistics to the last several years. Figure 3 shows the menu of different access statistics available. SciELO national coordination may also use Google Analytics to follow up the access.

Figure 3
Menu of downloads metrics at national collection with links to SciELO Brazil

Library collection

Library site usage reports

- Journal requests
- Articles requests by language
- Issue requests
- Top ten titles
- Articles per month

It is important to note that the SciELO interface provides access to the HTML texts and, from the HTML interface, access to the related PDF files that are stored as system files. The interface logs produce the download metrics that are available in the SciELO interface (Figure 3). However, the system PDF files that accompany the full-text HTML-based database are also independently available for indexing and harvesting by systems, indexers, and browsers such as Google. Therefore, in addition to access counts logged by the SciELO interface, users and robots directly download the journal articles in PDF format from the file system. According calculations in SciELO Brazil, access counts double when the direct access to PDF files is computed. As previously

mentioned, the development of a system capable of tracking access throughout the network and calculating related statistics, such as COUNTER, is of crucial importance for the SciELO Program.

The use of download metrics adds a new dimensions to the evaluation of journal and article impact. The indicators rank the humanities and social sciences journals higher than life and physical sciences. Unfortunately, there is no international, central control of downloads that allows for the comparability across journals published by different portals, which is why the metrics have been restricted to the SciELO universe. Another limitation SciELO faces with download metrics is the fact that open access articles are released in many institutional and thematic repositories whose downloads are not counted.

4.4. INTERNATIONAL SOURCES OF METRICS

Outside the SciELO Network, international tools and indexes, such as Altmetric.com, Google Analytics, Google Scholar, Web of Science (WoS), Scopus, and Webometrics provide citation, usage, and reference metrics on the performance of SciELO collections, journals, and articles. The major advantage that these metrics provide is the international comparability, while the major limitations are that WoS and Scopus do not cover all SciELO journals.

The Ranking Web of World repositories (Webometrics) operated by Spain's Consejo Superior de Investigaciones Científicas (CSIC) is a benchmark metric source on the online performance of the SciELO collections. While it is limited to open access portals, Webometrics shows how the performance of individual SciELO collections evolve across a group of ranking indicators such as visibility and number of documents. In the 2014 January edition, all SciELO Network national and thematic collections are present in the ranking of 143 top portals, with 8 of them in the upper quartile: Brazil (1st), Chile (8th), Argentina (11th), SciELO Public Health (15th), Spain (17th), Cuba (27th), Peru (34th), and South Africa (36th). SciELO Brazil has been ranked first since the creation of the Webometrics portal in 2008. The performance of the SciELO Network in Webometrics is good evidence that SciELO has accomplished its main objective to contribute to the visibility of the indexed journals.

Altmetric.com and similar services are able to account for the presence of articles and journals in online social networks, including references made on blogs, Facebook, Twitter, Mendeley, etc. SciELO established an agreement with Altmetric.com to collect this data for the articles published in the SciELO Brazil journals. In general, the presence of SciELO articles and journals online is, so far, very limited. The expectation is an increase in the coming 12 months when all

SciELO Brazil journals will have a minimum presence in the social networks. In July 2013, SciELO started to operate the collective blog *SciELO in Perspective* with a general topic dedicated to scholarly communication and another to the social sciences and humanities journals with the intention of adding a new topic in the near future. If the experience of SciELO Brazil with Altmetric.com is successful, other collections might also adhere to the contract.

Google Metrics is another international source that covers all SciELO Journals. Based on the Google Scholar database, it provides a journal's H-index metrics based on the citations received in the last 5 years in the ample and unrestricted universe of journals published on the Web. This universe favors the humanities and social sciences journals that performs better than life and physical sciences journals in correspondence with the downloads metrics. Google Metrics is also important as it allows the internal ranking of all SciELO Network journals, as well as some level of comparability with other journals.

WoS and Scopus are the main international multidisciplinary indexes used throughout the world to measure and rank scientific output based on number of articles and other types of documents according authors' geographic affiliation (countries, regions, states, etc.), authors' institution affiliations, researchers groups, and individual researchers. The presence and performance of SciELO journals in these indexes are indicators of visibility and quality based on compliance with their indexing policies. However, the SciELO, WoS, and Scopus indexing policies vary significantly. In the case of Brazil, only 25% of over 400 Brazilian journals indexed in 2014 are represented across all three indexes. SciELO Brazil has 70% of its journals also in Scopus, 37% in WoS, and 30% not present in either of these indexes. This discrepancy reduces the representativeness of these indexes. However, WoS and Scopus are the international benchmark indexes related to research impact and the relative performance of most of the SciELO journals are poor, as about 90% rank below the median distribution of the Impact Factor in their respective thematic areas.

The SciELO Citation Index (SciELO CI) was designed to overcome the limited coverage of WoS. It operates as one database index in the WoS platform and, therefore, permits the union of the SciELO Network and WoS journals. It represents a major effort of the SciELO Program to have an index covering a comprehensive collection of journals published at the national and international level that complies with a minimum set of selection criteria. It is possible to operate SciELO CI as a standalone database as well as combined

with other WoS database. In addition to improving the international visibility of journals and their research, SciELO CI will provide SciELO with a comprehensive universe of journals to follow and measure performance. The first ranking of the SciELO Network journals, which is based on the citations they receive within the universe of SciELO and other WoS Platform journals, will be published in late 2014. It will provide more comprehensive and robust performance indicators for both national and international oriented journals, which will contribute to research agencies and institutions as they work to enhance their policies on the development and evaluation of nationally published journals. SciELO CI will contribute to the follow-up of the internationalization of the journals based on the origin of the citations they receive, taking into consideration the thematic area and the language of publications. As previously acknowledged, most of the sciences and humanities journals, as well as a significant percentage of health sciences and agriculture journals, are highly nationally oriented.

5. CONCLUSION

Since its inception, the SciELO Program has prioritized the measurement of journal performance and the performance of collections of journals. In fact, as described before, from the very beginning the SciELO platform has encompassed an integrated performance reports module (Meneghini, 1998; Packer, 1998). However, the metrics currently used have strengths and limitations related to coverage and degree of comparability with non-SciELO journals. Although the classical bibliometric indicators are able to track SciELO journals performance over time at both the national and international level, they reveal a lower performance in relation to the journals from developed countries indexed in WoS and Scopus. The reasons for this lower performance are well known, starting with the fact that most of SciELO journals publish nationally-oriented research, which means research conducted predominantly by nationally affiliated authors and, in many cases, in languages other than English (Packer, 2014). Overall, the SciELO Network journals published over 80% of national authors and over 50% of articles in Spanish or Portuguese. The volume of domestic citations that these researchers may mobilize is not sufficient to compete with that of the international indexes. However, the evaluation of the journals by national agencies and institutions, and authors, privileges the high impact journals. The strategies SciELO is implementing to overcome this limitation is to operate with SciELO CI to maximize the capture of domestic and international citations and to promote the internationalization of journals to increase

the number of submissions from foreign authors. In this regard, it is also expected that Google Metrics will acquire more relevance in the future so its indicators are based on a wider universe of citations and serve as a reference source on performance, particularly for the social sciences and humanities disciplines.

Alternative metrics, such as downloads and presence in the Web, offer evidence of the performance of the SciELO collections and journals, but, as described above, they lack an international reference index that controls and certifies the downloads and therefore the comparability. For SciELO, these alternative and complementary metrics are of high importance since they measure a characteristic of SciELO journals and their usage by national communities of students, especially when nationally oriented research prevails. SciELO is adopting the COUNTER code of principles to calculate the metrics on downloads of articles. It is also promoting the usage of social network tools to improve the presence of the SciELO published research in the Web with the expectations of also using altmetrics as complementary metrics.

These strengths and weaknesses of the different metrics and their applicability to the different thematic areas, language of publication, and other specific characteristics of many of the SciELO journals pose a major challenge on the research and development of a multidimensional framework performance system to properly follow up its journals.

6. REFERENCES

BIREME / PAHO / WHO (2011). *VHL Guide 2011*, Version 19. São Paulo: BIREME/PAHO/WHO, 49p.

Meneghini R (1998). Avaliação da produção científica e o Projeto SciELO. *Ciência da Informação*, 1998, 27(2), p. 219-220.

MUGNAINI R, Tardelli AO, TUESTA EF (2004). Citations titles standardization using information retrieval techniques. In: *Journees internationales d'analyse statistique des donnees textuelles*, 7,2004, Louvain-la-Neuve. Actes JADT'2004 en ligne, 2004. p. 824-830. Disponível em: < http://lexicometrica.univ-paris3.fr/jadt/jadt2004/pdf/JADT_079.pdf >.

Packer AL, Biojone MR, Antonio I, Takenaka RM, García AP, Silva AC, Murasaki RT, Mylek C, Reis OC, Delbucio HCRF (1998). *SciELO: uma metodologia para publicação eletrônica*, 1998, 27(2), p.109-121.

Packer AL (2009). *The SciELO Open Acess: a gold way from the south*. Canadian Journal of High Education, 2009, 39(3), p111-126.

Packer AL (2014). *The emergence of journals of Brazil and scenarios for their future. Educação e Pesquisa*. 2014, 40(2), p. 301-323 .

SciELO (2013). *Action Lines for the Years 2014- 2016 with the Objective of Increasing the Visibility of the SciELO Network Journals and Collections*. São Paulo: SciELO, 21p

SciELO (2014). *Scientific Electronic Library Online,* www.scielo.org.

REDALYC

A PLATFORM OF VISIBILITY FOR THE SCIENTIFIC PRODUCTION

Published in Open Access Ibero-American journals*

Eduardo Aguado-López** and
Arianna Becerril-García***

1. INTRODUCTION

Today, the use of technological platforms that modify scientific production, communication, and practices of knowledge legitimization have gained greater relevance, particularly since the development of diverse regional initiatives that have been consolidating their on-line digital archives and libraries. These initiatives have had the goal of increasing the visibility of, and access to, the knowledge contained in the papers generated by universities and research centers whose purpose is to contribute to scientific debate, and who wish to do so in accordance to the disciplinary and social particularities of their respective national and regional agendas.

This phenomenon has not only oriented scientific production towards more democratic and inclusive communication scopes, but it

* Research Assistants: Miguel Leal Arriola, Salvador Chávez Ávila, Rodrigo Perera Ramos y Miguel Ángel Aguirre Pitol.
Editorial Review: José Luis Gómez Flores y Mauricio Pérez Sánchez.
Translation: Natalia Lifshitz

** Professor at the Autonomous University of the State of Mexico, Founder and General Director of Redalyc.org

*** Professor at the Autonomous University of the State of Mexico, Director of Systems and Information Technology in Redalyc.org.

has also invigorated the institutions themselves and their academic communities to the extent that they interact in more open and complementary directions. Moreover, we see the creation and integration of spaces with more collaboration among researchers, universities, and countries from different regions of the world.

This process was preceded by an indexing regime of academic publications that attempted to promote this universal dialogue, but which proved irresistible for the evaluation of the academic performance of the researchers and institutions. The alliance between the bibliometric databases and specialized editorial boards resulted in a restricted group of journals that were considered to be prestigious. This group of journals, these with access to indicators of their impact, have been those considered to be part of the "mainstream" in the communication and evaluation of science.

Thus, the alliance between academic journals,big publishers, and companies in charge of the dissemination of science started to consolidate.[1] This alliance generated a vicious circle in which researchers were looking for publications in venues higher in the hierarchies inside their disciplines in attempts to make themselves visible within this communication circuit—a circuit that placed emphasis on the need to increase the value of their citation indicators and, therefore, the measurable academic impact of their papers (Bourdieu, 1999).

Even though publishing in journals indexed in these databases may be seen as a source of credibility, a detailed analysis of the origin and topic of the papers according to the countries and the institutions of their authors can reveal little participation from countries in Latin America and the Caribbean – especially in the fields of social sciences and humanities. This is a reality that goes deep into each country because of their strong centralization of the scientific production centered in a few universities and some research centers (Russell & Ainswort, 2011).

At the same time, the idiomatic and thematic particularities that determine the diverse ways of production, communication, and collabo-

1 ISI-Thomson Reuters: Institute for Scientific Information was created in 1960, and has offered bibliography services, particularly citation analysis, since 1980. They produce an annual report Journal Citation Report (JCR) that uses the Impact Factor, that is, the mean number of citations of each paper in the journals it controls. The Ulrich index is a directory and database that provides information about periodical scientific publications. Its on-line counterpart, Ulrichsweb, has international coverage emphasizing publications in English. In turn, SciVerse-Scopus is an abstract and scientific journals citation bibliographical database, handled by the Dutch publisher Elsevier, which also offers author profiles based on affiliations, number of publications, and bibliographical data, such as the number of citations each published document has received.

ration of the science generated in the *global south*[2] often complicate the inclusion of the work of their researchers in the big databases, because they not only are linked with their communities' different habits and with the specificities of each source institution or country – particularly in Latin America and the Caribbean, regions strongly influenced by their deep social differences and inequalities – but they also account for a literary corpus differentially associated at international and local scales with the theoretical treatment and conceptual narratives of the theme-problems (Hicks, 2004).

In this sense, Chapter 4 of UNESCO's *World Social Science Report* (2010) shows how the number of social science articles generated between 1988 and 2007 included in the Social Science Citation Index (SSCI) of ISI-Thomson Reuters had its greatest increase in Latin America, despite the fact that the region is constantly underperforming Europe, North America, and Asia. This is why analyzing the Latin American production from open access regional platforms[3], such as SciELO, Latindex, CLACSO, or Redalyc, is particularly relevant as the region is more strongly represented in those scientific information systems (Beigel, s.f.).

Additionally, as mentioned by Dominique Babini (UNESCO, 2010), the main objectives of inter-institutional programs such as SciELO, CLACSO, and Redalyc are to increase the visibility and access of Ibero-American journals in order to develop regional indicators that allow for a more effective follow-up of scientific research, periodically providing diverse analysis on the progress and consolidation of regional networks. This is how they can function as a model for other organizations to develop similar or broader initiatives.

In that regard, it is important to remember the origin of the Network of Scientific Journals of Latin America, the Caribbean, Spain and Portugal (Red de Revistas Científicas de América Latina, El Caribe, España y Portugal), redalyc.org, which was founded in 2003 as an inter-institutional project at the Universidad Autónoma del Estado

---

2 This expression identifies what in other contexts and moments has also been referred to as "Developing Country" or "Third World". However, the idea of *global south* attempts to allude to the countries with medium and low income, generally located in the Southern Hemisphere, as opposed to Europe and North America. The expression is problematic in the case of Mexico, Central America, and the Caribbean, because even though they are located in the Northern Hemisphere, they share the same characteristics and similar problems as the countries located at the south of the globe, which is why we insist in using it as it allows us to refer to the countries that share problems related to low relative developmental levels, as well as particular organizational schemes which have allowed societies marked by injustice and economic inequality.

3 To identify the meaning of open access and its implications for scientific publications see Melero (2005) and Babini (2006).

de México (UAEMEX). Its objective is to create, design, and maintain an online library capable of compiling the peer-reviewed open access journals, as well as their scientific content, and to provide services for scientific information and act as a meeting point for everyone interested in consulting, debating. and validating the knowledge produced in Ibero-American nations. The mentioned, texts can be downloaded for free by students, academics, researchers, or any other interested person[4].

In addition to the aforementioned, redalyc.org has joined the websites that implement open access with high technological standards. Thus, interoperability mechanisms such as OAI-PMH (Open Archives Initiative – Protocol for Metadata Harvesting[5]) and exchange micro-formats encourage the linking and broadcasting of data from around the world. This situation allows for a greater dissemination of scientific information through a vast network that includes such highly relevant organizations as the Directory of Open Access Journals (DOAJ) of the Open Society Institute (OSI), JournalTocs, ScientificCommons.org, and Google Scholar, among others, and maximizes the access and impact of science produced in Latin America and the Caribbean at internationally competitive levels.

At the same time, the journals inside the redalyc.org project standout among the extensive scientific editorial production of Ibero America because they pass an evaluation process as a mandatory requirement. The criteria include international parameters of editorial quality, such as being ruled by peer review and the condition of publication, in their majority, original results from scientific research.

As a complement, the Scientometrics Lab redalyc-fractal (LabCrf) was created in 2010 as a research group in charge of analyzing the information associated with the papers in the database, with the goal of identifying and characterizing the behavior patterns of the science published in Ibero-American journals indexed by redalyc.org. One of the first concrete proposals of the LabCrf was the creation of

4 From its beginning, redalyc.org strove to bring together journals of social and human disciplines exclusively because, back then, they realized that those areas of knowledge were less likely to be incorporated into international databases or consolidated in their publishing processes. However, since 2006, the project opened up to the inclusion of journals from all areas of knowledge (Rogel-Salazar y Aguado-López, 2011).

5 The OAI-PM protocol outlines the generation of inter-effectiveness tools that, independently of the application, allow for the exchange of information so combined searches of the metadata of all the associated reservoirs (data suppliers) come from centralized points (service suppliers).

an analysis model based in scientific production and communication entities. For this, a set of scientometric indicators are applied whose goal is to account for the state of the art acquired by science produced inside the areas of knowledge and their disciplines at a regional, national, or institutional level – information of high utility for those who have to make decisions regarding scientific and technological development inside countries and institutions of the entities that participate in this database.

In that regard, initiatives such as redalyc.org gain great relevance because they efficiently increase the visibility and the interactivity around scientific papers throughout the internet, achieving a higher impact in the academic media and improving the communication among editors, readers, and authors. That is why in the *Latin American and Caribbean Consultation on Open Access to Scientific Information*, led by UNESCO in Jamaica in early 2013, it was recognized that the need to keep encouraging the work developed by the technological platforms and open access regional databases (UNESCO, 2013), particularly because all knowledge locked behind commercial barriers is sterile, as it remains confiscated by big publishing companies that take advantage of the research products generated with public funds (Llorens, 2013).

2. MAIN CHARACTERISTICS OF THE DATABASE

Redalyc.org is an online library that allows reading, downloading, and sharing of full text scientific papers for free, which is why it functions as a meeting point for everyone interested in reconstructing the scientific knowledge of and about Ibero America. This website – the most visible segment of this effort – is part of an initiative led by a group of researchers and editors concerned for the lack of visibility of the research results generated in and about the region. That is why it has been established as a window that allows observing the most noted scientific production in the Ibero-American region.

For more than 10 years, redalyc.org has been establishing itself as a relevant actor in the Ibero-American context, not just because it has been welcomed by the academic communities and their communication and scientific collaboration networks, but because of the its bigger contribution of providing a free and open content database. This includes the appropriateness of the information derived from the application of the scientometric indicators. These indicators, designed by the LabCrf with a novel and alternative focus, account for the behavior and the existing local and regional peculiarities around the process of generating scientific knowledge using public and transparent criteria.

In accordance with the latest technological and publishing developments, the appropriate strategy for scientific communication media edited in emergent countries is to be positioned so that it is relevant to the region using a combination of strategies aimed to improve quality and convenience, while also capable of capturing the best contributions of the researchers working on topics of interest to Ibero America and mainly to Latin America. These efforts will be able to contain the biases inherent to international collections whose tendencies have systematically affected the journals that are not seen as "mainstream," as mentioned by Ana María Cetto (Santillán, 2011). This is why we should move forward with actions that seek to reverse this import-export pattern of the science produced in countries from the *global south,* to be able to make an incursion in the validation and scientific debate from a more balanced position regarding First World countries (Guédon, 2011).

From this perspective, redalyc.org has functioned as a mediator that allows for the communication and indexing of the scientific production published in journals from the Ibero-American region through an open access technological platform, while also overcoming many obstacles to maintain and adapt the global standards to the work ways of scientists from the *global south* (Aguado-López et al., 2012). The Ibero-American contribution to the production of science was invisible because many of the communication spaces were absent in the databases that, from traditional models and standards, have certain access restrictions for countries, institutions, and researchers of the region. This is why the results presented next make what traditionally has been invisible, visible; that science has no borders.

Therefore, when taking into consideration that information technologies are in constant development and the process of science research and communication, it is relevant for knowledge, as a common good, to be available to whoever wants to access it. As such, redalyc.org is part of the recent movement of information exchange in open access technological platforms, which have experienced significant quantitative and qualitative growth in the last decade. This is why it contributes to the effort of strengthening the Ibero-American publications from editorial quality criteria, which improve the prestige of the journals and leverage the visibility of science generated in the region, emphasizing the work done by Latin American researchers in matters of humanistic and social interest.

3. INCLUSION AND EVALUATION OF JOURNALS

Today, scientific journals are not organizations of diffusion, but they also work as spaces that regulate the access to information and aca-

demic knowledge (Guédon, 2011). This is why they can be conceived as books in continual construction as continue publishing the daily work of researchers, who, apart from promoting the visibility of their contributions to the scientific field, also guarantee the quality of their academic hypothesis from the prestige some of the journals can offer.

In this context, redalyc.org builds its library in agreement with a series of policies and selection procedures of guaranteed quality journals. The journal collection is the raw material for the access and content retrieval services in the website and the data universe from where the metadata comes from, which informs the indicators applied by the LabCrf. Thus, the journals in this important project gather the editorial and academic standards. This is guaranteed by using an exhaustive methodology for adding them to the database that internationally accepted and validated parameters.

The evaluation methodology is made of 60 quantitative and qualitative criteria organized in three units – admission, quality, and management – with which the candidate journals are evaluated. These results are ratified by an international Advisory Scientific Committee.[6] In accordance with this, it is essential to favorably satisfy all the criteria from the first two areas (admission and quality) to continue with the evaluation, because they represent 39 criteria – and a subsequent score between 31% and 69%, from which a journal needs to achieve 82% to be included in the database. In turn, the third area foresees 21 points with qualitative value, which therefore does not influence the previous quantitative estimation.[7]

It is important to emphasize that a substantial part of the evaluation process is centered on the compliance of generally accepted standards of importance used to determine the scientific nature of a journal regarding editorial and content quality. These are the peer review double blind, the integration of an editorial committee, the originality of most of the published research results, and a regular publication schedule. This is stated by the basic admission criteria, as following:

---

6 Composed of academics with recognized international prestige as experts in the fields of natural sciences, social sciences, and the humanities, and with broad editorial experience. This referee organization is meant to support the evaluation guidelines to which journals will be subject to and to give their academic guarantee in the incorporation or rejection of any journal to the redalyc.org database. For a detailed review of the methodology, see: http://www.redalyc.org/info_pe.oa?page=/politica-editorial/metodologiaevalua.html

7 The total of criteria and the detailed evaluation methodology can be consulted in http://www.redalyc.org/info_pe.oa?page=/politica-editorial/metodologiaevalua.html

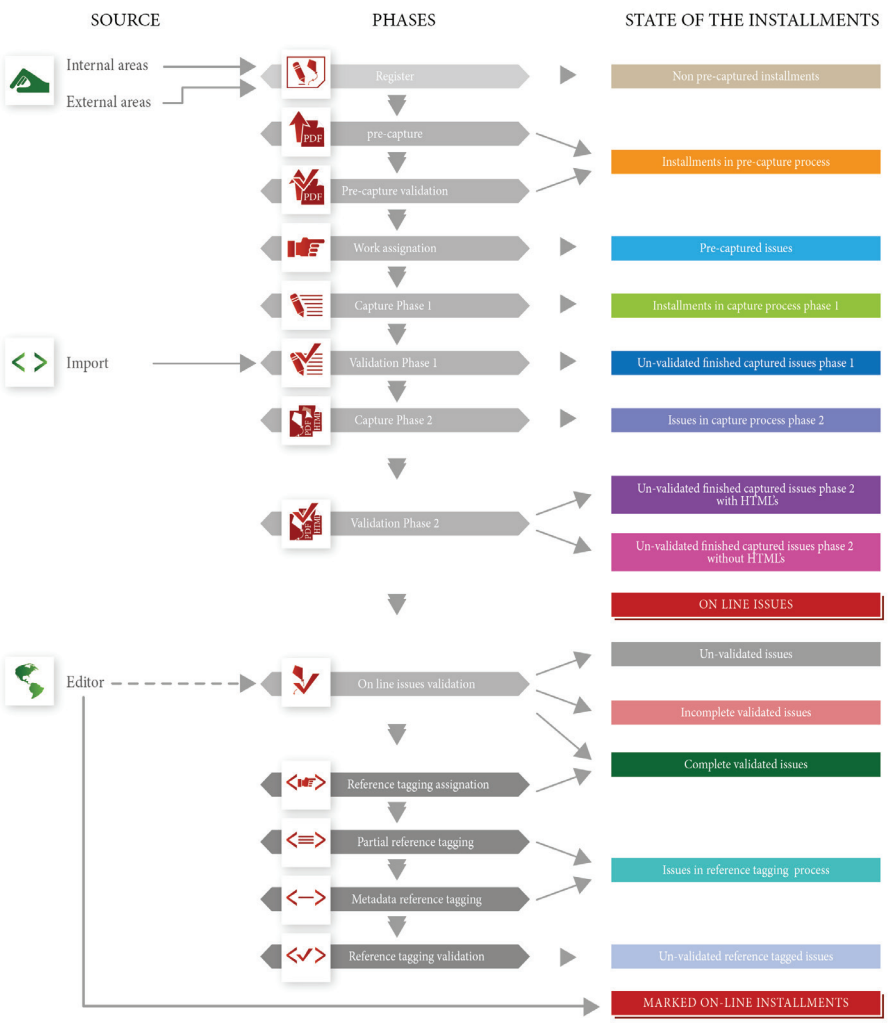
- Seniority and periodicity: the journals must have published constantly the year previous to their nomination, in order to prove seriousness and capacity to gather and publish their written materials.
- Originality: to prevent the science dialogue from becoming a monologue, it is recommended for papers to be submitted to only one journal, and to make original contributions to the theoretical and applied debate of the disciplines.
- Scientific content: the content of the journal must refer at least 50% to 75% to material from results of an academic research.
- ISSN: each publication must have a numeric code referring to the existence of a printed or electronic publication.
- Editorial committee: the journal must have a committee of experts on the subjects associated to the publication, whose members cannot be from the editing institution only.
- Appropriateness of the information: the papers must have title, abstract, and key words in the native language of the research, and an abstract in another language, in order for these contents to be considered by experts or readers from diverse latitudes.
- Referee system: in order to guarantee the quality and prestige, the papers must be reviewed by experts through the peer review, double blind system.

Adhering to these criteria, redalyc.org guarantees that the open access academic journals edited in paper or electronically are indexed in the database according to scientific and editorial quality standards. An indexing proof is given to every journal that joins the database.

4. REGISTRATION, VALIDATION, AND NORMALIZATION OF INFORMATION

To make the application of the scientometric indicators proposed by the LabCrf possible, the information associated with the scientific work published in journals of the redalyc.org database is uploaded to the Redalyc Information Integral System (RIIS). Through this system, the metadata identified in each paper are recognized, validated, and stored. This procedure allows for the association of each paper to one or many authors, each author to an institution (to where he/she has professional affiliation), and each institution to a country, according to the process detailed in Figure 1.

Figure 1
RIIS registration, validation and normalization process



This operation of identification and assignation, which might appear simple in its approach, faces the difficulties derived from the diversity and complexity of editorial practices. For example, there is no normativity to assign the institutional affiliation data of the authors and

most of the time each journal does this their own way. As such, the institutional affiliation identification depends upon an adequate reading of the author's signature, which can be: concise (when the author's name is followed by the name of an institution), complex (when the author's name is written along with a series of data that can be curricular or of affiliation), or invalid (when there is no affiliation data). That is why the determination of the country of the author's institution of affiliation also has similar difficulties and, likewise, cannot always be obtained.

Because of this, the registration of institutions sets out the challenge of identifying the ways in which each one of them can be assigned. For example, the Universidade Federal do Rio Grande do Sul can be identified by some authors and editors as UFRGS or as Universidad Federal del Río Grande del Sur, depending upon their norms, criteria or uses. That is why the RIIS allows a group of specialists to rely on evidence to make the decisions that allow the normalization of information and the identification of the different forms of institution nomenclature and to associate them with the official name. This way, it is possible to gather the production of an institution even though they appear in journal papers under different forms or languages.

It should also be noted that it is indispensable for each entity to be associated to a particular country, because there are homonym institutions that belong to different countries, such as the Universidad de los Andes, Colombia (Uniandes) and the Universidad de los Andes, Venezuela (ULA), or entities with many headquarters such as the Facultad Latinoamericana de Ciencias Sociales (FLACSO), with representation in Mexico, Chile, and Ecuador.

As stated above, the lack of normalization in the authors' names and their corresponding institutional affiliations makes hard to find the information in the database, because search engines generate disperse information according to the various ways an author or institution name is referred. This also affects the identification of citations and bibliographical references, and in consequence, the application and interpretation of indicators developed by the LabCrf to characterize the communication and scientific collaboration patterns.[8]

To settle these problems, in redalyc.org the data from institutions and authors' countries of affiliation are subject to the validation and

8 Even though the adequate identification of the institutions and countries in each paper can be considered a direct responsibility of the editors, who communicate with authors the importance of properly registering metadata, it is also the responsibility of the authors, who also incur the same problem when registering their own names with different variations.

normalization treatment described above, where the result is associated to a unique identifier corresponding to each author, institution, and country, according to the following sequence:

- Registration: capture of author information (full name and institutional affiliation) the same way the editor registered it in each paper of the journals in redalyc.org.
- Validation: it is verified that the institution is correctly registered and associated with the country specified in the paper.
- Normalization: a unique identifier is created for each institution form and the ones referring to the same organization are associated, in order to link the institutions detected as aliases and register their production in one unique entity.

This process is a continuous work and includes monitoring by qualified personnel who review the capture of information in different moments, because the new journals registered in the RIIS catalogue are incorporated with complete archives (from 2005 to date) and affect transversally the study years total.[9] In this sense, the information integrated in the database can be grouped in four categories: a) journal's general information (name, institution, country, area) and issue (number, volume, type of issue, publication's year and language); b) paper's data (title in original language, title in a second language, if possible, paper classification, abstracts, key words, received/accepted dates and first and last page); c) author's identification (name, last name, institution of affiliation, personal/institutional e-mail and author's signature – true copy of all the data recorded in the PDF file of the journal); and, d) institution's data (name, initials, institutional URL, street, postal code, country, sector and function).

This way, the entry and normalization information for each journal, paper, author, and institution allows the RIIS to systematize useful information to know how the communication and scientific collaboration between institutions, journals, areas, disciplines, and authors takes place:

- An indexing system that uses authorized and explicit criteria to incorporate journals into its database and guarantee its editorial quality.

9 The institution normalization process has been done for over 10,000 institutions registered in redalyc.org. A more detailed review would be the responsibility of expert documentalists from each country.

- A library with open access to the full text of on-line scientific papers, and with interoperable metadata to facilitate the location, visibility, and analysis.

5. COMPOSITION OF THE ON-LINE DIGITAL DATABASE

The redalyc.org database has more than 300,000 full text online contributions –growing by 4,000 a month. This material has been published by about 870 journals, which have been evaluated based on the double blind framework to prove its editorial and academic quality that guarantees the information has been previously academically reviewed.

Its thematic coverage is open to the most diverse scientific disciplines in science and the social and humane areas, and thereby integrates the academic journals published in any country from Latin America, the Caribbean, Spain and Portugal. Likewise, journals published in other countries can be integrated, provided that their topic of coverage is focused on problems related to Ibero America. These will be grouped under the heading: Latin-Americanists. Currently, scientific production from the social sciences represents 59% of the total disseminated material, with education, psychology and medicine the most productive disciplinary fields.

5.1. JOURNAL DISTRIBUTION BY COUNTRY

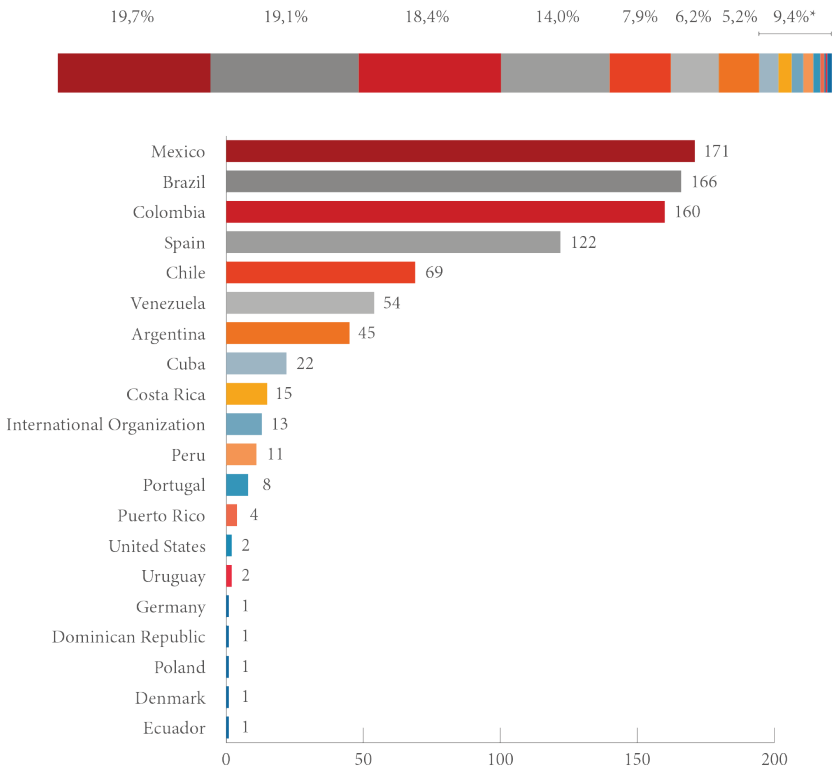
The following analysis corresponds to data collected in the first fortnight of October 2013, which consisted of 255,696 research papers that have been published in the database's 869 open access journals. Thus, being a database created with the purpose of giving visibility to the academic production of Ibero-American scientists, it is logical for Mexico, Brazil, Colombia, and Spain to be the countries with the highest contribution to indexed journals, with 71.2% of the total publications (see Graph 1).

Publications produced by Chile, Venezuela, and Argentina represent 19.3% of the total of publications registered in the database, followed by a smaller group of 12 countries and international organizations.

5.2. JOURNAL DISTRIBUTION BY AREA OF KNOWLEDGE AND DISCIPLINE

Regarding journal distribution by area of knowledge and discipline, it is important to mention that one of redalyc.org's main characteristics is the number of publications related to social sciences and the arts and humanities (514 and 95, respectively), which together represent 70% of the publications registered in the database (see Graph 2).

Graph 1
Redalyc.org journal distribution by country



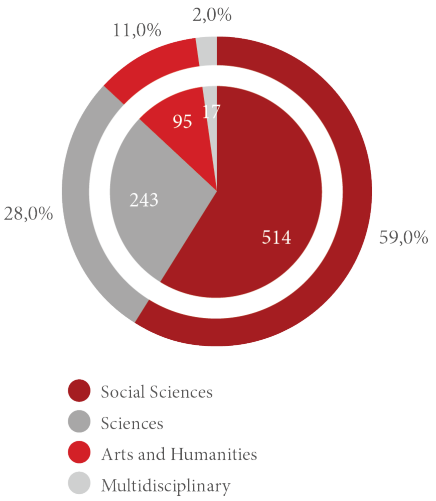
Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

* Includes: Cuba, Costa Rica, International Organization, Peru, Portugal, Puerto Rico, United Dominican Republic, Poland, Denmark and Ecuador.

Regarding the participation of publications according to disciplinary fields, it is important to highlight that education, psychology, medicine, sociology, and agricultural sciences are the first five top-producing fields, representing 36.9% of the journals in the database. However, it is very significant that even though the database has more journals in social sciences – probably because redalyc.org was originally focused in this study area –there are disciplines associated with exact and hard sciences that, in a shorter period, have gained enough importance to

be included in the database, such as medicine, agricultural sciences, engineering, and biology, which reflect 191 publications and 22% of the database. This also indicates that the database is also being considered by the (academic) science community, which has traditionally been better represented in other bases and bibliometrical indexes (see Graph 3).

Graph 2
Redalyc.org journal distribution by area of knowledge

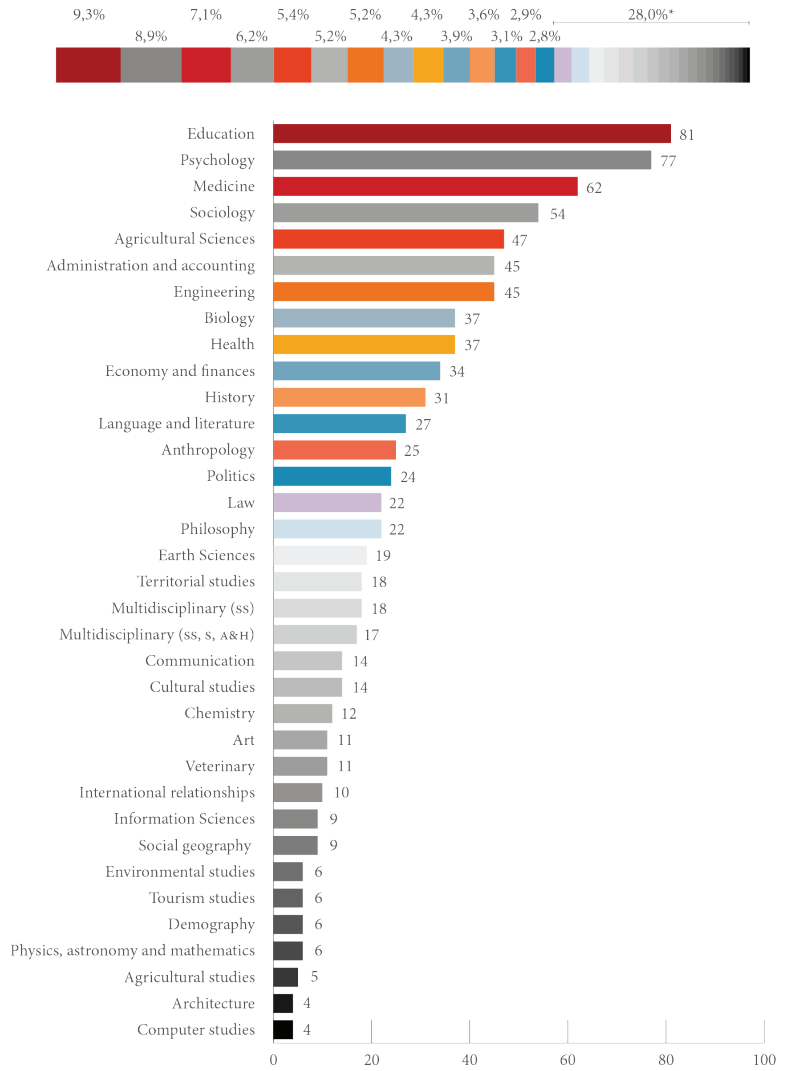


Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

5.3. PUBLISHED CONTRIBUTIONS BY KIND AND YEAR, AS WELL AS BY COUNTRY AND AREA OF KNOWLEDGE

For the 304,391 contributions stored in the dataset, it is possible to observe different levels of intensity throughout the publication of academic work. There are at least four major periods distinguishable when the data is viewed historically, among which the last two are the more representative with 18.6% and 77.7%, respectively. Additionally, from the total of work disseminated by the journals in the database, 83.9% correspond to scientific papers (255,696), with the remaining portion reflecting editorials, reviews, and other scholarly products. When looking at Graph 4, the huge effort redalyc.org has made to incorporate journals in the last 10 years is evident, as well as its contribution to the providing scientific papers available in full text.

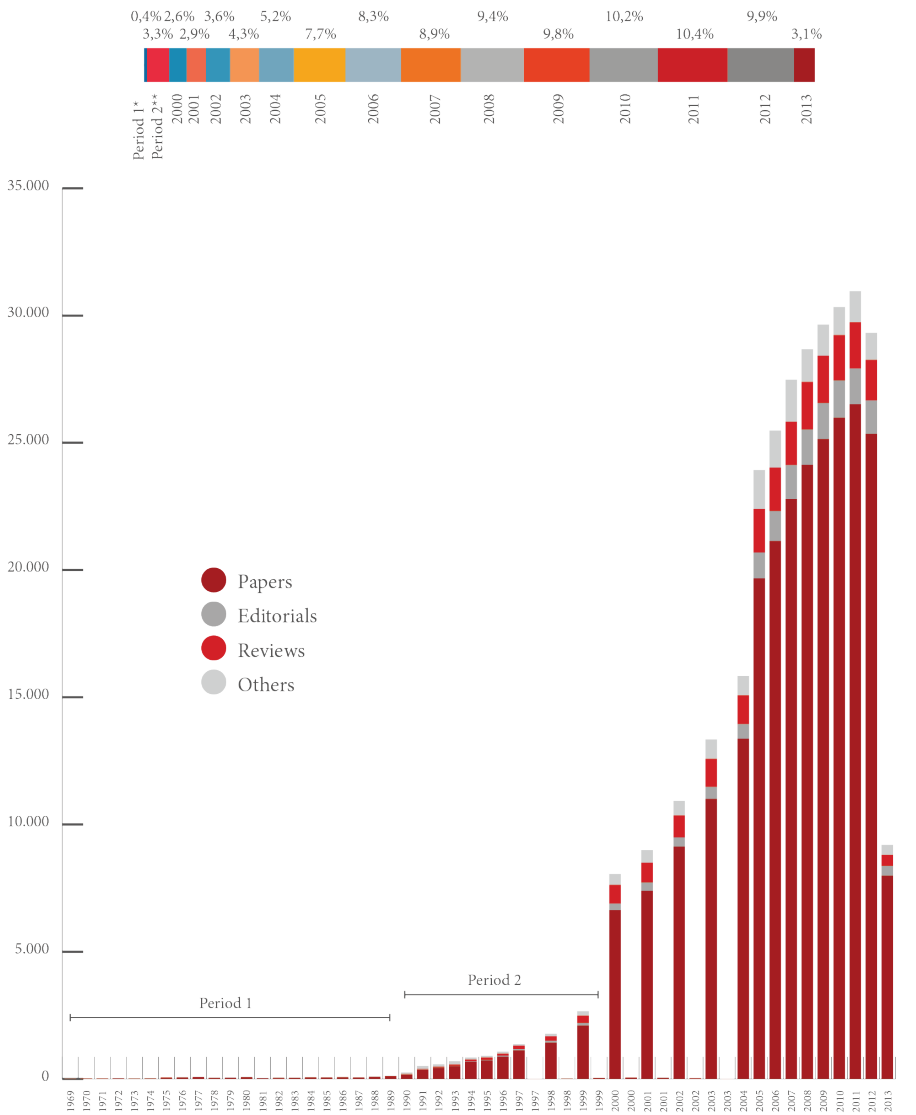
Graph 3
Redalyc.org journal distribution by discipline



Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

* Includes: law, philosophy, earth sciences, multidisciplinary (SS), territorial studies, multidisciplinary (S, SS, A&H), cultural studies, communication, chemistry, art, veterinary, international relationships, social geography, information sciences, enviromental studies, physics, astronomy and mathematics, tourism studies, demography, agricultura studies, architecture, computer studies.

Graph 4
Published contributions in redalyc.org's journals by year and kind

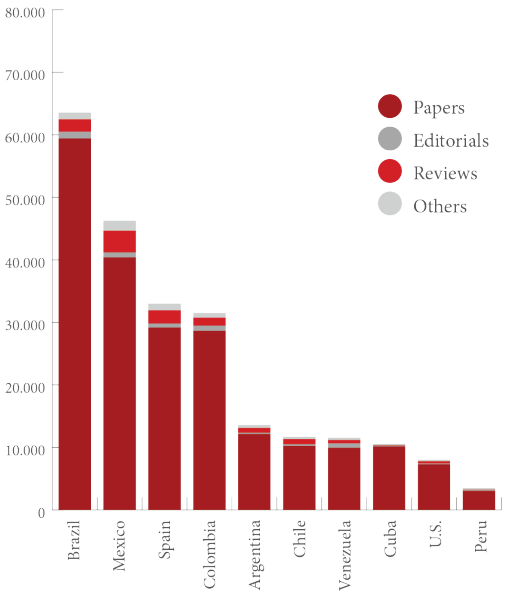


Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

* From 1969 to 1989
** From 1990 to 1999

When viewed by country (Graph 5), Brazil has the greatest contribution of academic work (more than 60,000 items), followed by Mexico (about 46,000), and Spain and Colombia (about 30,000 each). This is why the participation and behavior of these countries relates, in part, to the number of journals indexed by the database. It is important to point out that among the first 10 countries with highest contribution to the database, seven are Latin-American, while the remaining three are individually linked with the Caribbean, Spain, and North America.

Graph 5
Contributions published in redalyc.org's journals by kind and country

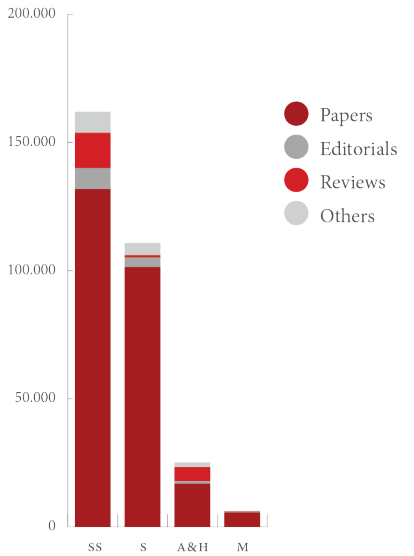


Country	Papers	Editorials	Reviews	Others
Brazil	59.406	1.060	2.012	1.070
Mexico	40.376	798	3.468	1.627
Spain	29.160	648	2.110	1.092
Colombia	28.643	839	1.252	745
Argentina	12.127	222	755	455
Chile	10.229	261	830	343
Venezuela	9.929	686	555	371
Cuba	10.129	116	84	241
U.S.	7.273	137	357	277
Peru	3.031	145	42	259

Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

The area with strongest academic support is social sciences, followed by sciences and, to a lesser extent, arts and humanities and the multidisciplinary field (Graph 6). Among them, the strong weight that scientific papers acquire becomes evident.

Graph 6
Contributions by kind and area of knowledge



Areas	Papers	Editorials	Reviews	Others
Social Sciences	131.840	8.205	13.727	8.263
Sciences	101.389	3.733	856	4.863
Arts and Humanities	16.816	996	5.538	1.857
Multidisciplinary	5.651	348	97	212

Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

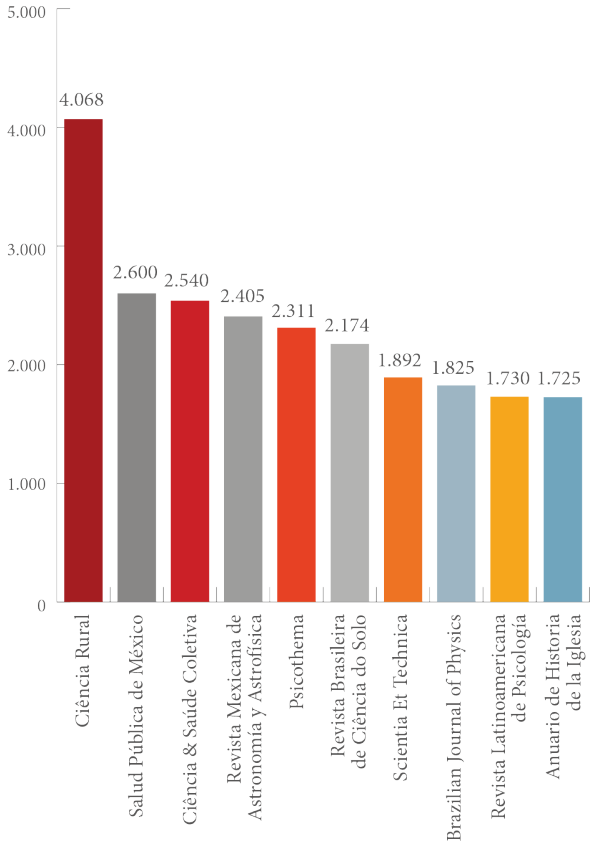
5.4. NUMBER OF PAPERS PER JOURNAL

The 10 journals that contribute more papers to the database are shown in Graph 7. Here, it is possible to see the balance between the two main areas of knowledge – five journals for sciences and five for social sciences – and the specific case of the journal *Ciencia Rural*, published in Brazil and focused on the discipline of agricultural sciences.

On a smaller scale, with 2,000-3,000 contributions, are also *Ciência & Saude Colectiva, Revista Mexicana de Astronomía y Astrofísica,*

Psicothema, and *Revista Brasileira de Ciência do Solo,* which represent the strong participation of Brazilian and Mexican publishers.

Graph 7
Papers by journal



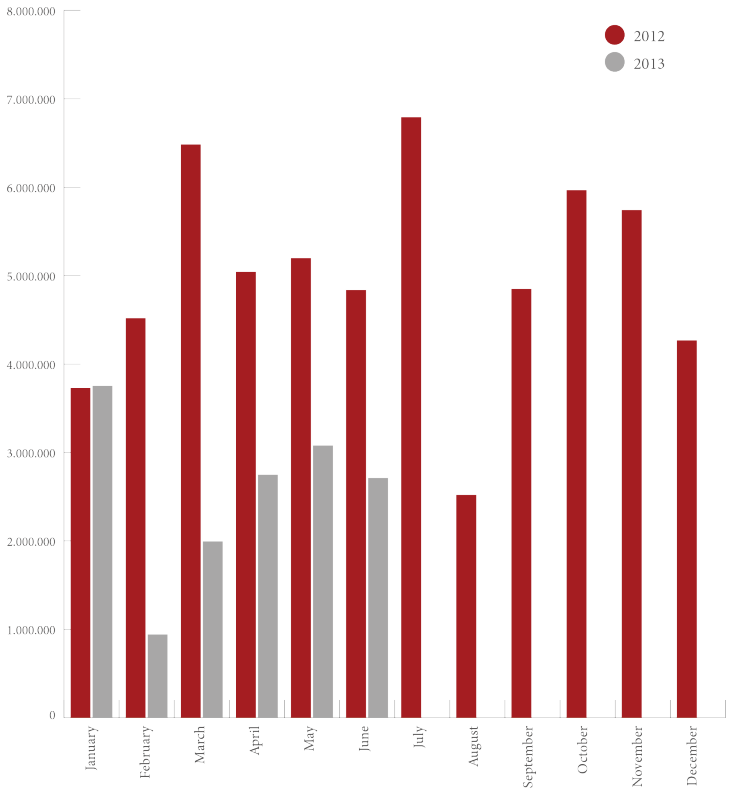
Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

5.5. USAGE STATISTICS

The following graphs show some initial data from the redalyc.org website on usage and the number of downloads for full text content accessed in different parts of the world. It is important to mention that the record filtering and georeferencing are performed using the list of robots included in the tool *awstats* for the generation of statistics.

Graph 8 shows the monthly downloads of redalyc.org texts in 2012. The fact that more than 79 million product downloads stands out, and speaks to the relevance and strong social impact generated by those databases that share scientific material for free to any interested audience. These types of open access policies contribute to a better informed society that is able to practice a more participatory, democratic ,and inclusive citizenship.

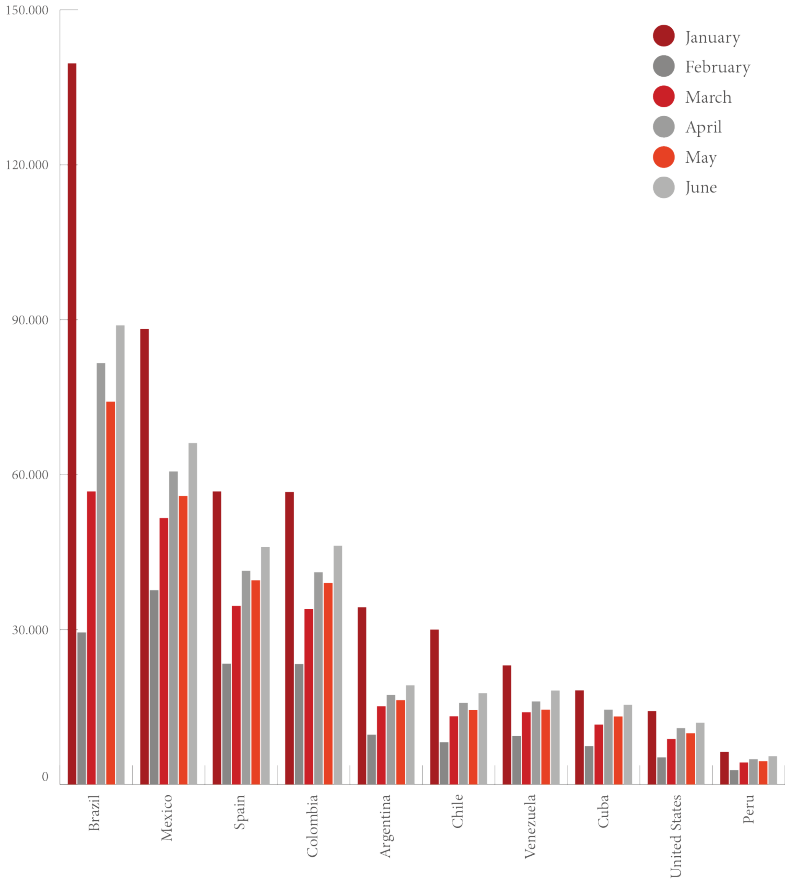
Graph 8
Monthly downloads of full text articles of redalyc.org 2012



Year	January	February	March	April	May	June	July	August	September	October	November	December
2012	3.729.468	4.517.791	6.482.745	5.043.012	5.197.375	4.836.733	6.790.604	2.520.631	4.849.685	5.966.959	5.741.458	4.267.697
2013	3.754.723	942.456	1.992.984	2.749.281	3.078.644	2.710.962						

Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

Graph 9
Monthly downloads of the redalyc.org content by country, first semester 2013



Country	January	February	March	April	May	June
Brazil	139.640	29.436	56.770	81.593	74.159	88.908
Mexico	88.216	37.642	51.610	60.623	55.867	66.121
Spain	56.772	23.385	34.605	41.393	39.562	45.979
Colombia	56.661	23.338	33.990	41.103	39.024	46.258
Argentina	34.329	9.675	15.150	17.338	16.347	19.206
Chile	30.010	8.211	13.228	15.830	14.449	17.689
Venezuela	23.098	9.414	13.996	16.110	14.513	18.217
Cuba	18.255	7.429	11.617	14.518	13.178	15.470
United States	14.223	5.255	8.825	10.921	9.925	11.986
Peru	6.334	2.795	4.308	4.896	4.529	5.473

Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

Brazil, Mexico, Spain, and Colombia are the countries that access the website the most and have the higher download rates of scientific papers during the first semester of 2013 (Graph 9). This participation confirms that a higher scientific dialogue is taking place amongst Ibero-American countries, not only related to communication and academic collaboration strategies, but to a higher consumption of scientific papers that are more appropriate and contextualized to the interests of the users and the database contents.

Although the United States is the only country external to the Ibero-American linguistic context that ranks in the top 10 countries with the highest redalyc.org downloads, there are many other countries in Central and Oriental Europe, Scandinavia, Africa, and Asia who frequently access materials published by journals in the database. This shows how science produced in the Ibero-American region is being consulted around the world, opening the path towards a higher internationalization of science from Ibero America and mainly from Latin America and the Caribbean.

Regarding article downloads by discipline (see Graph 10), it is interesting to see that among the 10 disciplines that are accessed the most, there is a balance in number of downloads between the areas of knowledge. This situation proves that the database is extremely comprehensive and does not present any kind of bias towards one area of knowledge or another. However, it should be highlighted that medicine and agricultural sciences are in the 1st and 2nd places in this regard, followed by education and psychology, and, to a lesser extent, sociology and health.

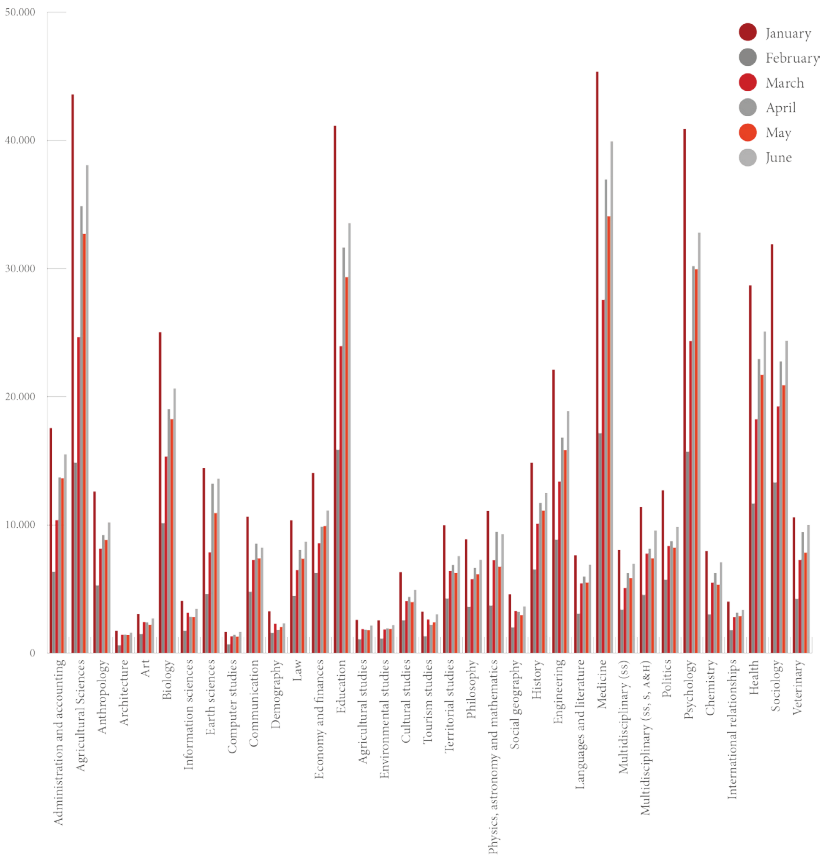
5.6. OVERLAP OF REDALYC.ORG JOURNAL DATABASE WITH OTHER DATABASES

As mentioned, journals from the Ibero-American region, particularly those from Latin America and the Caribbean, have encountered many barriers to inclusion in traditional international databases. In this regard, Sandra Miguel (2011) mentions the limited dissemination of Latin-American journals in international databases – mainly the Journal Citation Reports from Thomson Reuters, because Elsevier's Scopus has been broadening their inclusion and coverage policies towards Ibero-American journals. However, with the creation of regional initiatives such as SciELO and redalyc.org, the science produced in countries from the *global south* gained a closer tool to make its scholarly products more visible at regional and international levels.

These alternative databases are of great importance for the science produced in Ibero-American nations. There are 339 shared journals between redalyc.org and SCiELO published by 12 Ibero-American countries and two international organizations, and combined they cover more than 1,300 scientific journals published by an Ibero-American country that not

only publish internal and regional content, but also contributions from researchers affiliated to institutions all around the world (Graph 11).[10] This represents 36% of the total publications in SciElo and around 42% of the total in the redalyc.org database. The 5 countries that share more journals in both regional bases are: Colombia (90 journals), Mexico (81 journals), Brazil (57 journals), Chile (43 journals), and Argentina (26 journals) – with the first four countries alone representing almost 80% of the total of journals common to both online platforms (Aguirre-Pitol et al., 2013).

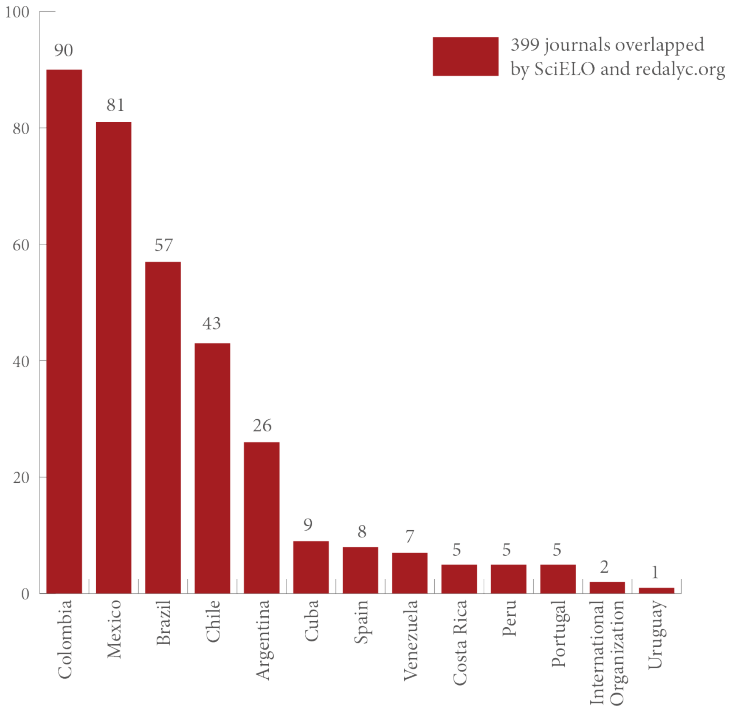
Graph 10
Monthly downloads of redalyc.org content by discipline, first semester 2013



Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

10 Its important to mention that this analysis was performed in February 2013.

Graph 11
Overlap of SciELO and Redalyc.org by country

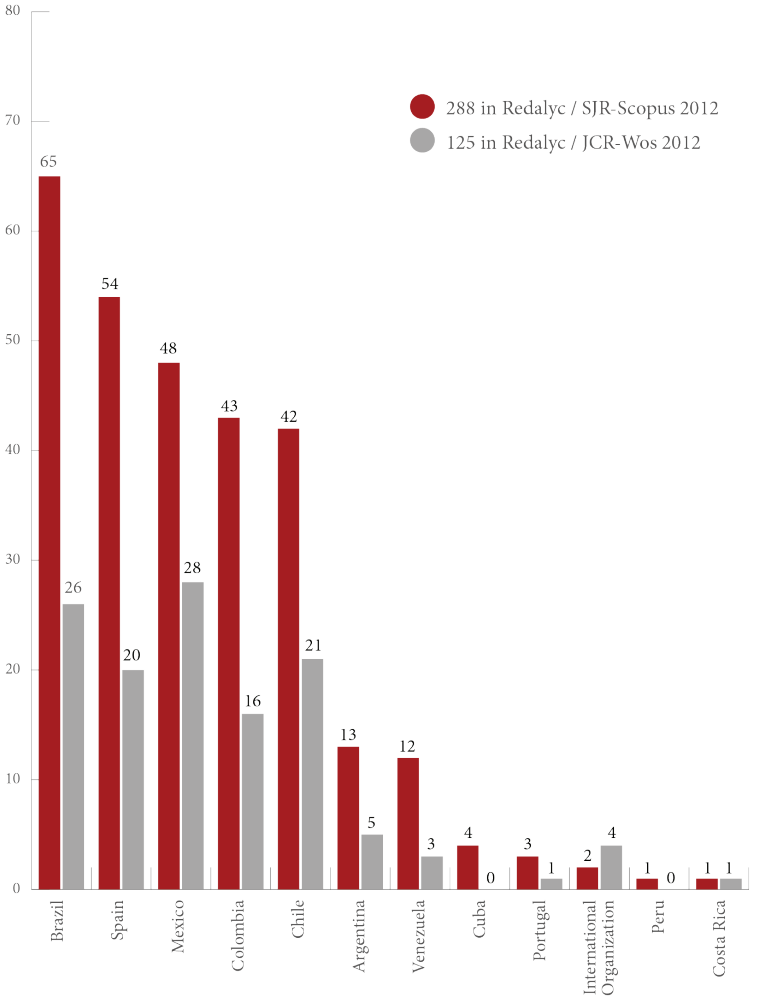


Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

Additionally, while SciElO and redalyc.org are databases that evaluate the editorial quality of open access journals, Thomson Reuters Web of Knowledge (through the Journal Citation Reports [JCR]) and Scopus-Elsevier (through their Scimago Journal Country & Rank index [SJR]) evaluate the bibliometric impact of the scientific production at a paper- and journal level. This is why each one systematizes different kind of data, apart from the goal of the information they produce and the geographical and thematic coverage of the scientific production they publish (Gasca-Pliego et al., 2013) (Graph 12).

As it is seen, the scientific journals from the Ibero-American region are underrepresented in "mainstream science." If we compare the redalyc.org database with the characteristics of this two indexes, we can see that they share just a few journals –288 for SJR and 125 for JCR – with most of them coming from Spain and Brazil.

Graph 12
Journals shared by JCR and redalyc.org, and by SJR and redalyc.org, 2012



Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. y http://www.scimagojr.com/journalrank.php y http://admin-apps.webofknowledge.com/JCR/JCR?SID=2AkMTXNMJG8bF9TZkfE| Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

The journals that do not overlap in these big databases of high academic prestige would be doomed to invisibility if not for regional projects such as redalyc.org. This is why this relevant initiative is a solid and consistent alternative that provides an alternative way to analyze the

generation of scientific knowledge in the Ibero-American context. It is worth noting that the bibliometrical bases built around the specialized publishers are having their first interactions with regional journals in an attempt to better the way in which the systems provide regional information, a decision that shows the higher relevance they have been acquiring for scientific communication (Vessuri et al., 2013).

Additionally, redalyc.org is ready to make an incursion into the field of *Almetrics* using social networks to learn about and measure the impact of the dissemination through these alternative metrics. This would allow authors to see the attention their scientific papers are receiving, while providing editors, librarians, and repository managers to assess the online activity around the academic literature they disseminate.

As stated above, the arguments pointing out that the content of regional databases are of low quality are misguided, because these journals can be found in Scopus, redalyc.org, and in SciElo. It would be relevant for the academic community and research groups to initiate a serious and responsible debate to distinguish, for the first time, the notions of quality and prestige that research journals may have, beyond the prejudice and false assumptions.

6. METHODOLOGY USED TO GENERATE INDICATORS

Since the scientific papers published in journals from the database are the center of analysis and research of the LabCrf, it allows us to determine the characteristics of the editorial capacity of the institutions and the countries of Ibero America, and to identify the elements needed to identify the different patterns of scholarly production, communication, collaboration, and usage of written science, in the Latin American region. We can observe, for example, how much of what is published is made public in journals from the same institution or country, what are the participation rates of foreign media and institutions, and the proportion and characteristics of scientific papers that are produced in collaboration with national and foreign academic peers (Becerril-García, 2012).

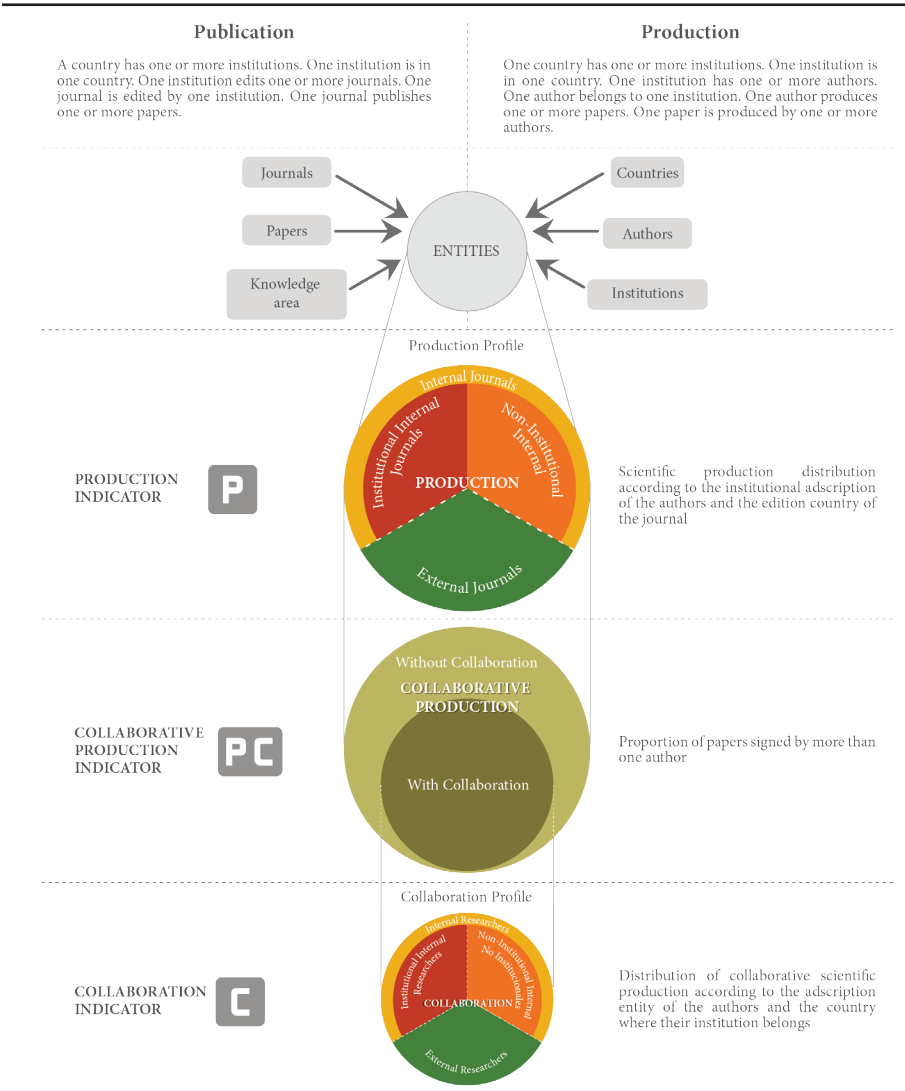
One of the objectives of the studies performed by the laboratory is to give information about the magnitude and possible impact of the strategies and practices adopted by countries, institutions, and researchers that contribute to the production of scientific knowledge over time. This kind of analysis captures the peculiarities in the communication and collaboration of scientific papers of a country or institution across journals in the redalyc.org database.

7. DESCRIPTION OF THE ENTITY-BASED ANALYSIS MODEL

To give greater clarity about the analysis model used below, a brief description of the interpretation criteria of the *Production and Col-*

laboration indicators and their *Internal-External and Institutional-Non Institutional* components is presented in Figure 2.

Figure 2
Interpretation of the indicators according to the entity-centered analysis model



Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

It is worth mentioning that the criteria used to interpret the composition of the indicators and their distribution always starts with where the entity is produced, because countries, institutions, and researchers are the only ones capable of generating scientific papers, while the area of knowledge, the disciplines, and the journals are the keepers of the produced work. In order to define *Internal-External and Institutional-Non Institutional* composition inside the *Production and Collaboration* indicators, it is necessary to turn to the corresponding level and to the producing entity, which is analyzed based on nits relationship to the where the entity is produced and the edition of the publishing journal.

8. INDICATORS OF PRODUCTION, COLLABORATION, AND USAGE

Once the main characteristics of the database are described, not only regarding their pertinence to the Ibero-American context but also in terms of the distribution of the journals and the corpus of papers that are a part of the studied universe, we then provide a description of the alternative indicators proposed by the LabCrf. These are a part of the scientometric studies set to explain the scientific production from a field broader than bibliometrics, because they are not restricted to the impact of the academic papers according to their level of citation inside journals included on international indexes, but they include the analysis of other determinant factors associated to the production of written science (Pérez Angón, 2006).

In accordance with the aforementioned, the process indicators developed from the analysis model were based on production and communication entities that the laboratory applied to papers published between 2005 and 2011 in some of the Ibero-American journals in redalyc.org. This data was used to generate a *Scientific Production Profile* determined by the characteristics and the behavior that each analyzed entity listed in the database. In this sense, LabCrf identified two main indicators obtained from entity metadata related to *Production (P)* and *Collaboration (C)*. These indicators allow for the identification of communication and collaborative work strategies used by researchers and institutions around written science from their components (*Internal-External and Institutional-Non Institutional*).

9. PRODUCTION INDICATORS

The *Production (P)* indicator is defined as the total number of papers produced by the analyzed entity and published in open access Ibero-American journals indexed by redalyc.org. Its construction is based on the relationship between the institutional affiliation of a paper's first author and the country of the entity that publishes the journal.

This indicator is composed by *External Production (EP)* and *Internal Production (IP)* according to the following terms:

- *External Production (EP)*. Integrated by the papers published by the researcher in a journal published by any institution from a country different from the country of its adscription entity. Additionally, due to the fact that the universe of journals is of Ibero-American origin, all the papers from researchers affiliated to non-Ibero-American institutions will be catalogued as foreign and, therefore, only the papers from researchers of institutions from this region can be classified as institutional and non-institutional internal information, other than external information published in any other Ibero-American country.
- *Internal Production (IP)*. Constituted by papers published by the researcher in a journal published by any institution located in the same country as its institutional adscription, which is subdivided by:
  - *Institutional Internal Production (IIP)*, constituted by papers published by the researcher in a journal published by the same institution where he researches and/or teaches; although, this can only be distinguished for institutions that have at least one indexed journal in the database.
  - *Non Institutional Internal Production (NIIP),* integrated by papers published by the author in a journal published by any institution other than its institutional affiliation, but located in the same country.

The components of the indicator (P) are clearly summarized in Table 1:

Table 1
Components of the Production (P) indicator

Internal Production	Links papers published in journals edited by institutions with the same country of affiliation as the author. This is subdivided by Institutional Production and Non-Institutional Production.
Institutional Internal Production	Relates papers published in journals edited by the same institution as the author's affiliation.
Non-Institutional Internal Production	Describes papers published in journals edited by an institution from the same country, but different to the author's adscription.
External Production	Refers to papers published in journals edited in a different country from the country of the author's institution of affiliation.

Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

This relationship is graphically shown in Figure 3:

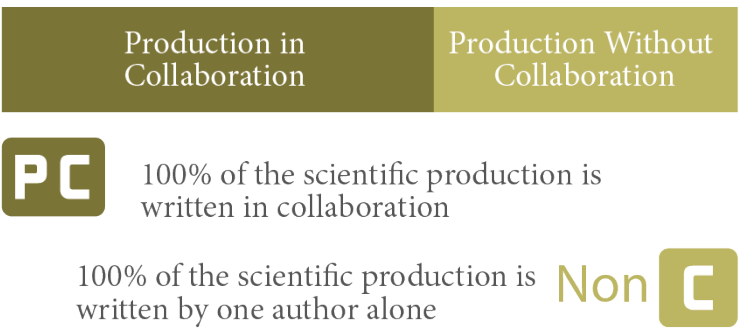
Figure 3
Distribution of the Production (P) indicator

Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

10. INDICATORS OF COLLABORATION

The *Collaboration (C)* indicator is based on co-authorship and is limited to the group of papers that, related to the total production, are written by a minimum of two researchers from any region of the world who decide to communicate their work together in an open access Ibero-American journal of redalyc.org. Papers written by one author are classified as *Without Collaboration (WC)*. This distribution is shown in Figure 4.

Figure 4
Distribution of the Collaboration (C) indicator



Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

As stated above, this is obtained from the relationship between the institutional the country of the entity and the institutional affiliations of the researchers

participating in co-authorship such that when more than two countries or more than two institutions appear, it is possible to analyze the patterns showing the *External Collaboration (EC)* and the *Internal Collaboration (IC)* as detailed below:

- *External Collaboration (EC)*. Composed of papers authored by two researchers from any region of the world, where the external character depends upon the participation of at least two countries whose institutions have one or more researchers contributing to the co-authorship of the scientific paper.
- *Internal Collaboration (IC)*. Established by the papers in co-authorship among researchers whose institutional affiliation are located in the same country. This is subdivided in:
  - *Institutional Internal Collaboration (IIC)*. Composed of papers written by at least two researchers affiliated to the same academic or research institution.
  - *Non Institutional Internal Collaboration (NIIC)*. Composed of papers created by at least two researchers affiliated to different institutions located in the same country.

Conceived this way, the components of indicator C are clearly shown in Table 2:

Table 2
Components of the Collaboration (C) indicator

Internal Collaboration	Refers to contributions written in collaboration exclusively by authors from the same country. The Internal Collaboration is subdivided by: Institutional Internal and Non-Institutional Internal.
Institutional Internal Collaboration	Links papers written in collaboration exclusively between authors affiliated to the same institution.
Non-Institutional Internal Collaboration	Relates papers written by authors affiliated to different institutions from the same country.
External Collaboration	Describes papers published in collaboration with authors affiliated to one or more institutions of the analyzed country, with authors affiliated to institutions of countries different from the analyzed country.

Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

Additionally, the relationship among components of this indicator appears in Figure 5:

Figure 5
Distribution of the Collaboration (C) indicator

Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

11. DATA UNIVERSE FOR THE CALCULATION OF METRICS

The analysis performed on the database in October 2013 used as a data source the set of 800 open access journals indexed by redalyc.org, which have published 145,515 research papers between 2005 and 2011. From now on these will be denominated as *Paper core* (see Table 3). Although the database had more than 800 journals during the period of this study, only the titles with complete online content with analyzable metadata were considered.[11]

Table 3
Data universe of analysis for the application of redalyc.org metrics 2005-2011

Source Universe	Total
Analyzed journals	800
Countries that register scientific production	146
Paper core (scientific production)	145,515
In collaboration	95,263
Without collaboration	50,252
Institutions with scientific production	13,414
With contribution in social sciences	7,181
With contribution in sciences	8,413
With contribution in arts and humanities	1,311
With contribution in multidisciplinary	1,066
Scientific production by continent	153,318
Scientific production by country	156,734
Scientific production by institution	206,335

Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

11 A journal is considered to have *complete content* when all its issues are available in electronic format through redalyc.org, in function of the declared periodicity.

The study only considered the research papers and essays published between 2005 and 2011, which altogether represent 90.1% of all the academic contributions published in journals indexed in the database. Because of this, contributions such as editorials, presentations, reviews, and various texts were not considered for the scientometric analysis, as shown in Table 4.

Table 4
Contributions analyzed for the application of indicators
in the database redalyc.org, 2005-2011

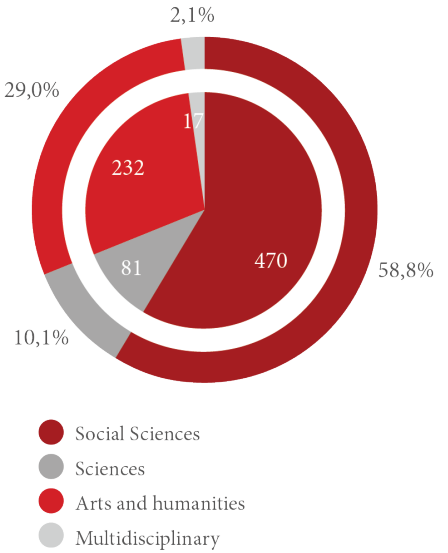
Types of Contributions	Absolutes	Relatives
Papers and/or essays	145,515	90.1%
Editorial and/or presentation	3,491	2.2%
Reviews	8,171	5.0%
Other documents	4,263	2.7%
Total	*161,440*	*100.0%*

Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

From this set of papers, 95,263 were written in collaboration, which means that more than half of the analyzed production (65.5%) came from a work in co-authorship involving two or more researchers that may have the same nationality and belong to one institution, or may have different nationalities and/or belong to different institutions. Such papers constitute the basis to explain the characteristics of the collaboration around scientific output, where it is possible to develop statistics by country, and the type of the coauthors' institution of affiliation (see Table 1).

Regarding the distribution by area of knowledge and discipline, one of the main features of redalyc.org lies in the number of journals participating in social sciences, arts and humanities; these journals represent 68.9% of all the publications in the database (see Graph 1), followed next by publications in the fields of education, psychology, and sociology, which represent 23.6% of publications. Additionally, it is appropriate to highlight the speed with which the database has been embraced by the academic community in science, particularly in the field of medicine, agricultural sciences, and engineering (18.2% of the journals). This composition is shown in more detail in Graphs 13 and 14.

Graph 13
Distribution of the source journals for the calculation of metrics by area of knowledge, 2005-2011

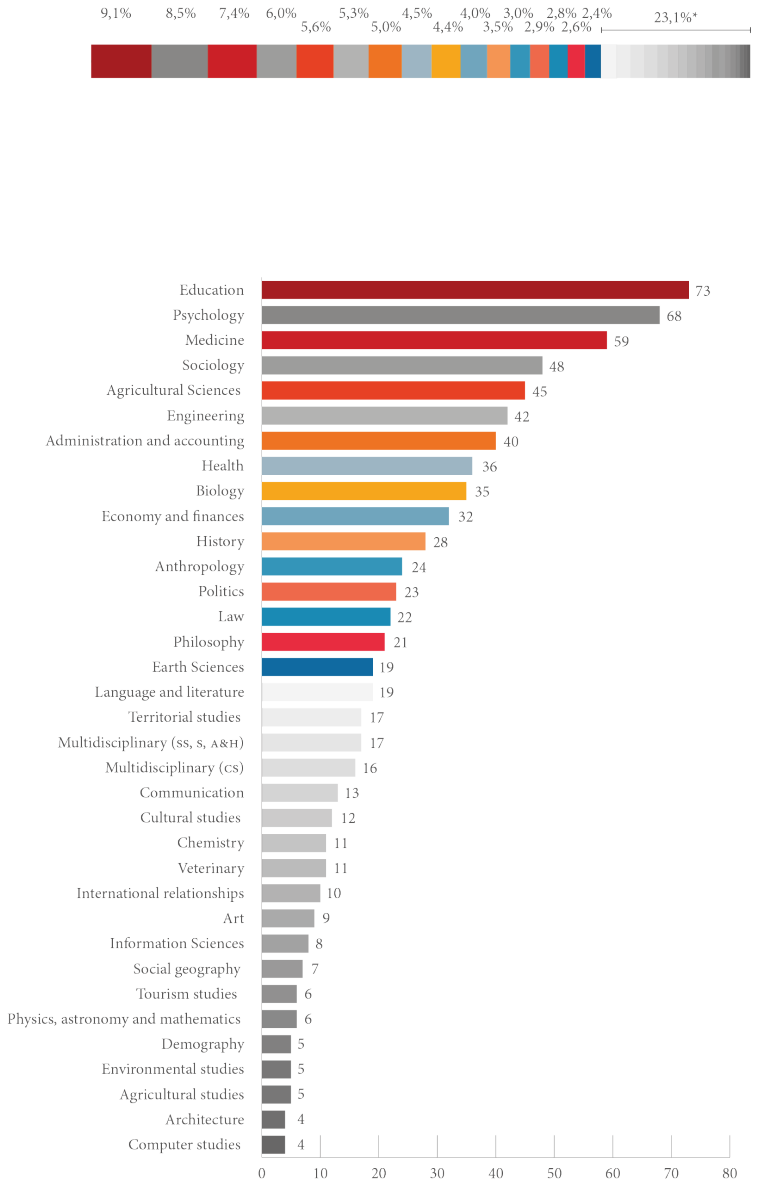


Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

Excepting international organizations[12], 15 countries publish the journals indexed by redalyc.org and participate in the calculation of the indicators (see Graph 15). When analyzing the scientific output based on of the country of the authors' institutional adscription, it must be noted that the number of nations that publish their research results in journals of the database increases to 146 and, in different magnitudes, it covers countries from all continents.

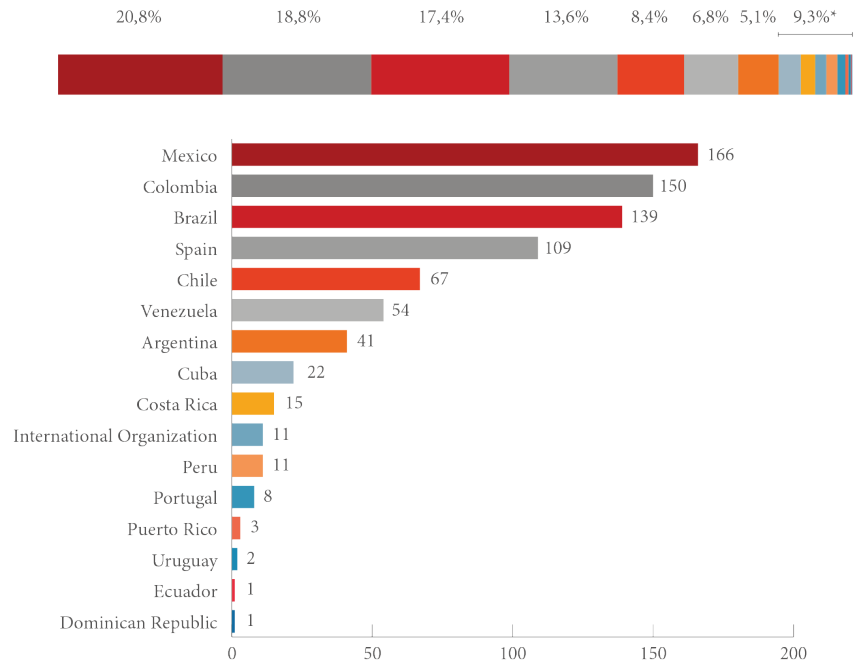
12 Some examples of international organizations that stand out by their amount of contributions in redalyc.org are: the Sociedad Interamericana de Psicología with 333 papers, the Centre International de Recherches et d'Information sur l'Economie Publique, Sociale et Coopérative with 221 papers, and the Institut Français d'Études Andines with 179 papers.

Graph 14
Distribution of source journals for the calculation of metrics by discipline



Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

Graph 15
Distribution of source journals for the calculation of metrics by country of edition, 2005-2011



Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

* Includes: Cuba, Costa Rica, International Organization, Peru, Portugal, Puerto Rico, Uruguay, Ecuador and Dominican Republic.

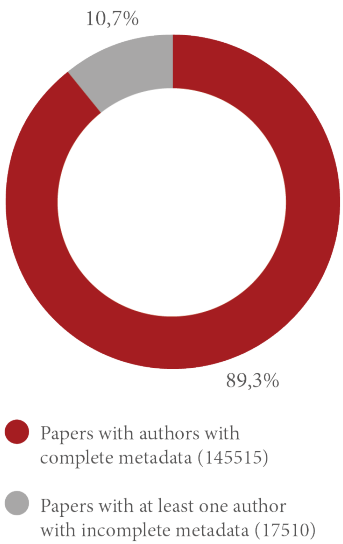
At the same time, the total number of institutions with papers published in any of the redalyc.org journals between 2005 and 2011 was 13,414. Among these, 8,413 were from the sciences; 7,181 from the social sciences; 1,311 from arts and humanities; and 1,066 from multidisciplinary fields, as shown in Table 1. This composition exhibits a relatively balanced distribution between the number of institutions that participate in sciences and social sciences. This is a good example of the input of the Ibero-American institutions in the communication of scientific knowledge around the region, compared with the disciplines of the "mainstream science".[13]

13 For a better explanation of the "mainstream science" in relation to peripheral science, see Guedón, 2011.

To know the magnitude of the scientific output by country and authors' institutional affiliation – and considering that one paper can be authored by more than one author– the *core of papers* was disaggregated so one paper can be considered as many times as different countries and/or institutions sign it. This is a very important because if affects the total amount of *Production* by Continent, Country, and Institution mentioned in Table 1, coverting the core of papers into 153,318, 156,734 and 206,335, respectively.

It is important to mention that even though the author's information is included in the papers, they do not always offer data on institutional affiliation, or such information is not specified enough or the institution's country is not mentioned either. These cases are considered as *authors with incomplete metadata*. Graph 16 shows the composition of the studied universe in this regard.

Graph 16
Authors with complete and incomplete metadata, 2005-2011



Source: Elaboration: scientometrics Lab Redalyc-Fractal (LabCrf) | Data from redalyc.org. | Metodology: http://www.redalycfractal.org/met | Creation: october 18th 2013.

The *authors with incomplete metadata* are excluded from the analysis, so the patterns of authored by two authors, but one of them has in-

complete metadata, the paper will be considered as published by only one author and, as such, classified as *without collaborati*on, which means that paper will not count for the institution and/or country of the author with incomplete data.[14]

In conclusion, using redalyc.org as a source of scientometric analysis for research papers allows for the following:

- It is the information system with the highest number of Ibero-American journals with complete contents from 2005 to 2011, which makes it a significant database of high utility to perform different kinds of analysis.
- From the Ibero-American websites, it is the database with more journals in Spanish.
- It is the information system with the lowest levels of bias and with one of the best processes of metadata registration, validation, and normalization.
- Almost two-thirds of the database is represented by social sciences, arts, and humanities journals. This describes the editorial universe that defines the database's strengths and representation. However, this is balanced with the area of sciences regarding the number of scientific papers due to the higher periodicity and publishing rate in journals of this area of knowledge.
- It has a set of criteria to guarantee the journal quality. These fulfill international scientific quality parameters, and are supervised by an International Scientific Advisory Board that includes renowned researchers, each one related to different areas of knowledge, disciplines, and lines of research.
- To be included in the online journal database, the system requires the fulfillment of international standards of editorial quality, the existence of electronic files of all the papers published from 2005 to date, and the acceptance – through a goodwill agreement – of the open access model.

14 We are aware of the implications of this methodological decision. However, it is based on the following logic: a) less than 5% of the total source universe belongs to this category, and b) including papers with incomplete metadata in the analysis would have necessitated adding the category "unidentified" to the country and the institution. This would mislead the reader because in the majority of the cases, the problem is not the author's lack of institution or country of affiliation; the problem is the set of omissions due to editorial care. By eliminating these cases from the analysis, we ask for the minimum responsibility of every editorial quality process and, by making this clear, we make the implications of incurring in bad editorial practices, visible.

12. FINAL CONSIDERATIONS

As stated, the fact that journals from countries in the *global south* are marginally represented in the "mainstream science" databases, where many results and research subjects go unnoticed, is unquestionable. This is not just a consequence of authors' publishing in their native language, it is also because in the special case of social sciences, most of the research and the papers that account for it are strongly linked to local interests and problems. Their research, then, may not necessarily reflect the established approach to publication from an international perspective, which contributes to why their research is often invisible to the systems of specialized publication and their existent measuring tools (Gingras & Mosbah-Natanson, 2011).

As mentioned by Alperin (en Adams, n.d.), the majority of the Latin American journals are mainly published by public academic institutions, which is why instead of looking for profit, they want to create communication spaces to encourage a greater dialogue inside and outside their academic communities. In this sense, the redalyc.org database has the social role of supporting public universities and their editorial projects – often times subject to diverse restrictions that impair a greater dissemination of the knowledge they produce – to the extent that it gives a set of tools to their academics and researchers, and to their editorial teams, which effectively increases the visibility and the interactivity around the scientific papers through an extensive network of usage and collaboration via the internet.

This online library has its similarities with the simple yet powerful inspiring force of the great Library of Alexandria – the capability of gathering the greatest amount of publically-funded knowledge produced and provided by a diversity of cultures with the noble goal of sharing it with anybody who is interested, without any economical, technical, social, or legal restrictions (Aguado-López, 2013).

However, today, all the scientific disciplines are involved in a generalized race towards the evaluation of their impact from bibliometrical indicators that take into account the number of citations one way or another. This has wrongly been seen as a measure of their quality, and in the specific case of social sciences and humanities, it has played against the diversity and critical character of their academic advances. Therefore, this way of measuring, generally out of context, has proven to be ineffective for the humanistic and social disciplines, especially for the knowledge produced outside the theoretical and methodological paradigms accepted by the mainstream that are often not published in the journals classified in the rankings of private companies such as the Reed-Elsevier and Thomson Reuters (Vessuri et al., 2013).

This transformational process has opened up academic groups and their research communities beyond localized dialogue and regional collaboration, pushing the Latin American sciences towards the internationalization of their scientific work strategies in such a way that once their analytical construct has been debated and agreed, they try to disseminate the knowledge in different ways and through different collaborations and publication fields, mainly through scientific papers that capture in a more efficient way the actual state of publishing as a highly changeable social phenomena.

The relevance of studying research output, collaboration, and communication from Ibero America, especially in Latin America and the Caribbean, resides in characterizing each of these elements inside the process of science production according to the particularities of each country, with the objective of providing a more appropriate approximation about what happens in this region, which has historically been analyzed from more ethnocentric points of view. That is why it is important to qualitatively and quantitatively document the way science is moving from local to global – from the identification of a set of communication and collaboration networks integrated at a regional and international level, to *glocal*[15] themed problems, which are often are associated with the millennium's development objectives established by UNESCO, such as poverty, migration, access to education and health, or climate change, to give some examples.

In this context, the ways for measuring scientific research in Ibero America are being incrementally debated, even when the available information is fragmented and difficult to give a precise and accepted diagnosis in the matter (Buquet, 2013). In this sense, although the mainstream databases keep leading the indicators that weight the investigative capacity at a global scale[16], and even though the bibliometrical rankings are recognized to help unify the organizational field of science (Sanz & de Moya-Anegón, 2010), these can be dangerous if

15 The term *glocal*, proposed by Roland Robertson comes from the interaction between global and local perspectives when building a culture that is global but has distinctive local characteristics that make it unique.

16 The Impact Factor (IF) is the mean of the amount of citations by the number of published papers in the journals considered in the annual JCR (Journal Citation Report) report of the ISI; the H index is presented as an alternative to the IF and implies a mean between the number of publications and citations they receive. Finally, the Relative Impact (RI) is applied by discipline and is the result of the quotient of the impact presented by a certain discipline in a country, divided by the impact of that discipline in the world according to the total of citations divided by the total of papers exclusive of that area.

used to formulate generalizations, which is why they must be qualitatively and quantitatively treated from a different perspective (Archambault & Lariviére, 2010).

This way, papers published in the journals indexed by redalyc.org and the metadata systematized by the LabCrf, fulfill the purpose of making the invisible, visible. It is important to remember that, from an epistemological perspective, no community can account for of any knowledge that is not made visible if it is not published and if society will not read it (López-López, 2010).

In this sense, redalyc.org is a highly significant tool for those who design scientific policies inside nations and Ibero-American institutions, and for those responsible for implementing them in the scientific field, including academics and researchers interested in these matters. This has made redalyc.org worthy of many awards and acknowledgments from institutions as important as the Instituto de Información Científica y Tecnológica (IDICT), Cuba; the Red de Investigadores sobre Globalización y Territorio RII, Belo Horizonte, Brazil; the Consejo Superior de Investigaciones Científicas, Spain; the Universidad Rey Juan Carlos and the Sociedad Latina de Comunicación Social; Universidad Complutense de Madrid, Spain; the Universidad de los Andes, Venezuela; and the World Summit Award.

Additionally, it must be recognized that redalyc.org strongly encourages the continuous improvement of the editorial processes of the journals that are part of the database, as well as for those that will be added in the future, according to standards of scientific quality that respect institutional, national, and regional peculiarities of each of the participating entities.

It should be pointed out that redalyc.org web portal has supported the Budapest Declaration in Latin America and the Caribbean, not only through the encouragement for publishing in open access journals, but also through the motivation for a Mexican legislation that stimulates the open access to scientific information and the existence of institutional repositories that allow better preservation and dissemination of research papers among the academic community and citizens interested in knowing the results of the research mainly developed with public funds.

As stated by Ordorika (2012), the country must continue legislating about public resources for production and dissemination of scientific knowledge, using technologies so public universities can keep broadening the sphere of their social responsibility without losing their essence as institutional formers of knowledge. That is why redalyc.org has been consolidated at a regional level while also having an international presence.

Therefore, this initiative invites institutions and their research groups to make this data a subject of study, capable of being analyzed longitudinally and across countries, institutions, areas of knowledge, and in relation to other technological platforms and similar online databases.

Finally, through the use of metrics calculated by databases with a significant coverage, such as redalyc.org, the details of the research results communicated in journals today and published by the Ibero-American region can be known. This contributes to the debate around the spaces and policies to which those in the center and and around the periphery of science pay attention to. Thus, thanks to the Scientific Production Profiles that can be obtained with the information from database, and to the resulting analysis by the academic community, redalyc.org and the Scientometrics Lab presents an extensive solution to understanding the contributions made by countries, institutions, and authors in this open access database, no matter their size, resources, infrastructure, or age. Behind this great effort, lays a central objective whose goal is to contribute to making the invisible, visible, because what can not be seen does not exist and scientific knowledge must be a common good available to everyone.

13. REFERENCES

Adams, C. (n.d.). Open access in Latin America: Embraced as key to visibility of research. *SPARC*. Retrieved from http://www.sparc.arl.org/news/open-access-latin-america-embraced-key-visibility-research-outputs

Aguado-López, E., Fischman, G., y Babini, D. (2013). ¿Son los promotores del Acceso Abierto los alejandrinos del Siglo XXI? In *Cuadernos del Pensamiento Crítico Latinoamericano* (CLACSO), *6*. Retrieved from http://biblioteca.clacso.edu.ar/clacso/se/20131016025858/Cuaderno-No6-SegEpoca.pdf

Aguado-López, E., Garduño-Oropeza, G., Rogel-Salazar, R., y Zúñiga-Roca, M. F. (2012). The need and viability of a mediation index in Latin American scientific production and publication. The case of the Redalyc System of Scientific Information. In *Aslib Proceedings* (Bradford) Nº 64. Retrieved from http://www.redalyc.org/redalyc/media/redalyc_n/acerca-de/inc/doc/Documento03.pdf

Aguirre-Pitol, M. A.; Leal-Arriola, M., y Martínez-Domínguez, N. D. (2013). Análisis comparativo de la cobertura de SciELO y Redalyc.org. México, LabCfr®-UAEMEX. *Repositorio*

Institucional UAEMEX (Toluca, Mexico). Retrieved from http://ri.uaemex.mx/handle/123456789/571

Archambault, É., y Larivière, V. (2011). Los límites de la bibliometría en el análisis de la literatura en ciencias sociales y humanidades. In *Informe sobre las Ciencias Sociales en el mundo. Las brechas del conocimiento* (Paris: UNESCO). Retrieved from http://www.foroconsultivo.org.mx/libros_editados/informe_sobre_las_ciencias_sociales_en_el_mundo.pdf

Babini, D. (2006). Acceso abierto a la producción de ciencias sociales de América Latina y el Caribe: bibliotecas virtuales, redes de bibliotecas virtuales y portales. *Dominique Babini y Jorge Fraga (comps.) Edición electrónica, bibliotecas virtuales y portales para las ciencias sociales en América Latina y el Caribe* (Buenos Aires: CLACSO). Retrieved from http://bibliotecavirtual.clacso.org.ar/ar/libros/secret/babini/Babini.pdf

Becerril-García, A., Aguado-López, E., Rogel-Salazar, R., Garduño-Oropeza, G. y Zúñiga-Roca, M. F. (2012). De un modelo centrado en la revista a un modelo centrado en entidades: la publicación y producción científica en la nueva plataforma redalyc.org. In *Aula Abierta* (Oviedo: Universidad de Oviedo), *40*(2). Retrieved from http://dialnet.unirioja.es/servlet/fichero_articulo?codigo=3920933&orden=0

Beigel, F. (n.d.). Latin American publishing system in the crossroads of the world-academic system. In *Sage/International Sociological Association, Current Sociology* (in process).

Bourdieu, P. (2012). Dos imperialismos de lo universal. In *Intelectuales, política y poder* (Buenos Aires: Eudeba/Clave intelectual).

Buquet, D. (2013). Producción e impacto de las ciencias sociales en América Latina. In *Doc. de trabajo del Consejo Latino Americano de Ciencias Sociales* (Buenos Aires: CLACSO). Retrieved from http://biblioteca.clacso.edu.ar/clacso/becas/20131016084109/Buquet_Ciencias_Sociales_America_Latina.pdf

Gasca Pliego, E., Martínez Carbajal, R., Tinoco García, I., Arriaga Jordán, C., González López, S. y Aguado López, E. (2013). *Informe sobre la producción científica de la UAEMEX en revistas iberoamericanas de acceso abierto en Redalyc.org 2005-2011* (Toluca, Mexico: UAEM-Facultad de Ciencias Políticas y Sociales Laboratorio de Cienciometría Redalyc-Fractal).

Gingras, Y. y Mosbah-Natanson, S. (2011). ¿Dónde se producen las ciencias sociales? In *Informe sobre las Ciencias Sociales en el*

mundo. Las brechas del conocimiento (Paris: UNESCO, Foro Consultivo Científico y Tecnológico). Retrieved from http://www.foroconsultivo.org.mx/libros_editados/informe_sobre_las_ciencias_sociales_en_el_mundo.pdf

Guédon, J.C. (2011). El acceso abierto y la división entre ciencia principal y periférica. In *Revista Latinoamericana de Ciencias Sociales* (Buenos Aires: CLACSO), *6*.

Hernes, G. (2011). Prefacio. Un planeta, mundos aparte ¿el mismo mapa? In *Informe sobre las Ciencias Sociales en el mundo. Las brechas del conocimiento* (Paris: UNESCO, Foro Consultivo Científico y Tecnológico). Retrieved from http://www.foroconsultivo.org.mx/li- bros_editados/informe_sobre_las_ciencias_sociales_en_el_mundo.pdf

Hicks, D. M. (2004). The four literatures of social science. In H. Moed (Ed.), *Handbook of Quantitative Science and Technology Research* (Netherlands: Kluwer Academic). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.10.500&rep=rep1&-type=pdf

Llorens, A. (2013). Le Monde: la importancia del acceso abierto y el ejemplo latinoamericano. Retrieved from http://editarenmexico.wordpress.com/2013/04/01/le-monde-la-importancia-del-acceso-abierto-y-el-ejemplo-latinoamericano/

López López, W. (2010). Internacionalización y visibilidad del conocimiento. *Universitas Psychologica* (Colombia: Pontificia Universidad Javeriana), *9*(2). Retrieved from http://www.redalyc.org/articulo.oa?id=64716832001

Melero, R. (2005). Acceso abierto a las publicaciones científicas: definición, recursos, copyright e impacto. In *El Profesional de la Información* (Barcelona: Swets Blackwell), *15*(4). Retrieved from http://eprints.rclis.org/6571/1/EPI-rmelero.pdf

Miguel, S. (2011). Revistas y producción científica de América Latina y el Caribe: su visibilidad en SciElO, RedALyC y SCOPUS. *Revista Interamericana de Bibliotecología*, *34*(2).

Ordorika Sacristán, I. (2012). Tareas pendientes de la política en educación superior y la importancia de reafirmar la visión progresista de la autonomía universitaria en el contexto actual. *Perfiles Educativos* (Mexico), *34*. Retrieved from http://132.248.192.201/seccion/perfiles/2012/nea2012/mx.peredu.2012.ne.p176-180.pdf

Pérez Angón, M. A. (2006). Usos y abusos de la cienciometría. *Revista Cinvestav* (México: Cinvestav). Retrieved from http://www.

cinvestav.mx/Portals/0/SiteDocs/Sec_Difusion/Revista-Cinvestav/enero-marzo2006/6%20usos.pdf

Quevedo-Blasco, R. y López-López, W. (2011). Situación de las revistas iberoamericanas de psicología en el Journal Citation Reports de 2010. *Universitas Psychologica* (Bogotá), *10*(3). Retrieved from http://www.redalyc.org/articulo.oa?id=64722377023

Rogel-Salazar, R. y Aguado-López, E. (2011). Redalyc ocho años haciendo presente la ciencia iberoamericana en el contexto de la Sociedad de la Información. In *María de Lourdes López-Gutiérrez, José Luis López-Aguirre & José Samuel Martínez-López (coords.) La comunicación que necesitamos, el país que queremos*, (Mexico: XV Encuentro Nacional CONEICC). Retrieved from http://es.scribd.com/doc/71612437/Ebook-La-comunicacion-que-necesitamos-El-pais-que-queremos

Russell, J. M., y Ainsworth, S. (2011). Investigación en ciencias sociales en América Latina en comparación con China e India. In *Informe sobre las Ciencias Sociales en el mundo. Las brechas del conocimiento* (Paris: UNESCO, Foro Consultivo Científico y Tecnológico). http://www.foroconsultivo.org.mx/libros_editados/informe_sobre_las_ciencias_sociales_en_el_mundo.pdf

Santillán-Aldana, J. (2011). La necesidad de sistemas de información abiertos no excluyentes: Entrevista a Ana María Cetto Kramis. *Biblos* (Peru), *42*. Retrieved from http://www.redalyc.org/articulo.oa?id=16120044005

Sanz Menéndez, L. y de Moya Anegón, F. (2011). Una nueva industria: la calificación de las universidades en las ciencias sociales. In *Informe sobre las Ciencias Sociales en el mundo. Las brechas del conocimiento* (Paris: UNESCO, Foro Consultivo Científico y Tecnológico). Retrieved from http://www.foroconsultivo.org.mx/libros_editados/informe_sobre_las_ciencias_sociales_en_el_mundo.pdf

UNESCO. (2010). World social science report 2010: knowledge divides. Paris, France: Author. Retrieved from http://www.unesco.org/new/en/social-and-human-sciences/resources/reports/world-social-science-report-2010/

UNESCO. (2013). Report of the regional Latin American and Caribbean consultation on open access to scientific information and research-concept and policies. Retrieved from http://www.unesco.org/new/en/communication-and-information/resources/news-and-in-focus-articles/all-news/

news/first_regional_latin_american_and_caribbean_consultation_on_open_access_to_scientific_information_and_research

Vessuri, H., Guédon, J.C., y Cetto, A. M. (2013). Excellence or quality? Impact of the current competition regime on science and scientific publishing in Latin America and its implications for development. *Current Sociology*. http://csi.sagepub.com/content/early/2013/12/02/0011392113512839

OPEN ACCESS INDICATORS IN SUBJECT DIGITAL REPOSITORIES

THE CASE OF CLACSO´S LATIN AMERICA AND THE CARIBBEAN SOCIAL SCIENCES DIGITAL REPOSITORY INDICATORS

Dominique Babini*, Gustavo Archuby**,
Valeria Carrizo*, Darío A. García*,
Sebastián Higa* and Diego Spano***

THIS LATIN AMERICAN COUNCIL of Social Sciences (CLACSO) report describes and analyzes indicators available for the monitoring and evaluation of the Digital Repository that provides open access to research output from CLACSO´s academic network.

CLACSO[1] is a non-governmental organization created in 1967 and currently encompassing 371 research centers and 654 graduate social science programs in 23 countries across Latin America and the Caribbean (also referred to in this appendix as "the Region"). Its aims are the promotion and development of research and teaching of social sciences; the strengthening of exchanges and cooperation among institutions and researchers within and outside the Region; and, the appropriate dissemination of the knowledge generated by social scientists, within the academic environment, as well as among social actors and movements, civil society organizations, and the general public within the Region. CLACSO also seeks to provide international visibility to

* CLACSO, Latin American Council of Social Sciences biblioteca@clacso.edu.ar.
** National University of La Plata (UNLP), Argentina (gustavo.archuby@gmail.com).
*** Prodigio Consultants (diegospano@gmail.com).

1 www.clacso.edu.ar

the Latin American and Caribbean social sciences, a production that achieves scarce visibility in international indexing services (Alperin, Fischman, & Willinksy, 2011; Archambault & Larivière, 2010; Cetto, Alonso-Gamboa, & Córdoba González, 2008).

To achieve the aims of dissemination, visibility, and access to research results, as of 1998 the CLACSO library, together with the editorial department, launched a network of digital libraries. In 2002, CLACSO started a social science Digital Repository[2] for Latin America and the Caribbean in an open source Greenstone[3] platform recommended by UNESCO. Since 1998, this program was conceived as a virtual space and online community for information exchange, experimentation, and cooperation in the development of open access scholarly communications throughout CLACSO´s network and in Latin America and the Caribbean in general. Priority is given to transfer knowledge to communities of practice where publishers, libraries, academic, and multimedia staff from member institutions participate, to strengthen their knowledge about open access. Also, the program contributes to building capacities that allow implementation of open access initiatives in the Region, improvements in the quality of the open access publications, and their dissemination in collaborative and institutional digital repositories in the Region. CLACSO´s Campaign in Support of Open Access to Knowledge holds activities for the promotion of open access research. This gives CLACSO the opportunity for active participation in national, regional, and international debates on open access alternatives and policies.

The specific objective of CLACSO's Digital Repository is to provide greater visibility and access to the research results produced by the Council members. Currently, the Digital Repository provides open and free access to a collection of 33,040 full texts journal articles, books, working documents, theses, conference papers, and a peer-review journal collection with 63 full-text journals of CLACSO in Redalyc. The multimedia collection adds 406 audio and video productions produced by CLACSO members.

Starting in 2009, a self-archiving system was developed that allows CLACSO member institutions to deposit metadata of digital objects (e.g., books, book chapters, working papers, theses, articles, journals, videos, audio, multimedia, etc.) thru a web-form with 15 fields according to Dublin Core[4] standards, which allow description of the digital object which is attached, except for media where a link is pro-

2 http://biblioteca.clacso.edu.ar

3 Greenstone is an open source software for digital repositories www.greenstone.org

4 Dublin Core Metadata are used to describe digital objects http://dublincore.org/

vided to the server where it is hosted. Within the form, each field has a minimal explanation with its corresponding examples. Once the text and its metadata are self-archived, a quality control process begins of the completed form, the attached files, and thematic and geographical descriptors. Each approved deposit is available for consultation in the Digital Repository's Reading Room – an online space organized in collections for each Member Center and CLACSO program.

This service is a decentralized, collaborative work -supported by CLACSO Members, by the Swedish International Development Cooperation Agency (SIDA) and the Norwegian Agency for Development Cooperation (NORAD)- and undertaken together with a community of publishers, libraries, and the multimedia community, with 1,055 participants from Latin America and the Caribbean.

Since 2009, CLACSO has an agreement with Redalyc who manages and funds the peer-review journal collection. Redalyc also has the responsibility of evaluating the journals that apply to be included in CLACSO´s collection, provides technical support and training of publishers, processes the journals and their metadata to add them to the collection of CLACSO journals in Redalyc[5] (currently 63 in total), and to the Redalyc collection in general.[6]

1. INDICATORS FOR THE EVALUATION OF THE DIGITAL REPOSITORY OF CLACSO

Since the beginning of the Repository, CLACSO defined that the monitoring and evaluation of the Repository would be made from three perspectives:

- *Community:* The extent to which CLACSO manages to promote open access, including the publishers and libraries of in its network and virtual community, and regularly expose them to the trends of academic communications in open access, contributing to a cultural change in the Region and the adoption of open access for the dissemination of research results.
- *Collections:* How the digital collections of the Member Centers and programs of CLACSO´s Digital Repository – including the collection of CLACSO journals in Redalyc – grow, taking care of geographical and institutional diversity, and performing specific actions to promote participation with contents in open access digital repositories from the Region.

5 http://clacso.redalyc.org

6 http://redalyc.org/home.oa

- *Use:* How the Repository and its digital collections are used: number of downloads, countries of origin of requests, most requested digital objects, subjects most consulted. Starting in 2014, metrics at article level is experimented with some collections.

CLACSO's Digital Repository website offers, in open access, the statistical reports as they develop. These reports allow the tracking of the above-mentioned three variables of community, collection, and use of the Repository.

2. INDICATORS OF COMMUNITY

CLACSO has invited publishers and libraries in its network of research centers to join a virtual community and participate in the exchange of information about international trends for scientific communications in open access and news about initiatives in the Region. The evaluation of this activity is performed by verifying the incorporation in the community of staff from the publishing, journal, library, and (sometimes) academic areas of new centers that are incorporated in CLACSO´s network; and, by checking the delivery of fortnightly professional updating communications from CLACSO to the community.

This virtual community is regularly informed about trends in open access scholarly communications, promoting good practices, quality formats, and the dissemination of research outputs in interoperable digital libraries/repositories according to updated international standards. To prepare these communications, the program performs monitoring and research of international and regional trends on open access and digital repositories for the selection of good practices, news, and articles to share with the community.

At the end of 2013, CLACSO's community of publishers and libraries[7] had 954 participants from 21 countries in the Region.

Considering the growing use of multimedia support for the dissemination of research results, in recent years CLACSO´s Digital Repository has started a Multimedia Exchange Community[8] to contribute to the sharing of mutual knowledge and the exchange of information between those involved in productions of audio, video, and other multimedia products in CLACSO´s network. It also promotes the incorporation of multimedia contents and its corresponding metadata in digital repositories. At the end of 2013, this community had 86 members from 14 countries. Furthermore, the Digital Repository

7 http://biblioteca.clacso.edu.ar/somos/equipo/comunidad/editoresybiblioteca

8 http://biblioteca.clacso.edu.ar/somos/equipo/comunidad/multimedia

now contains a multimedia collection with the metadata of 406 audio and video records of CLACSO´s network. Recently, the Executive Secretariat of CLACSO has launched CLACSO-TV[9], a program for web dissemination of short interviews, documentaries, and audiovisual records with the most representative personalities of the Latin American social sciences and culture, and with guests from other regions.

3. INDICATORS OF COLLECTIONS

The main digital collections in CLASCO's Repository are detailed below.

• **The Reading Room:** This online "reading room" (Figure A1) is organized by country and, within each country, by collections of each Member Center. The Reading Room runs on a platform developed with Greenstone open software.

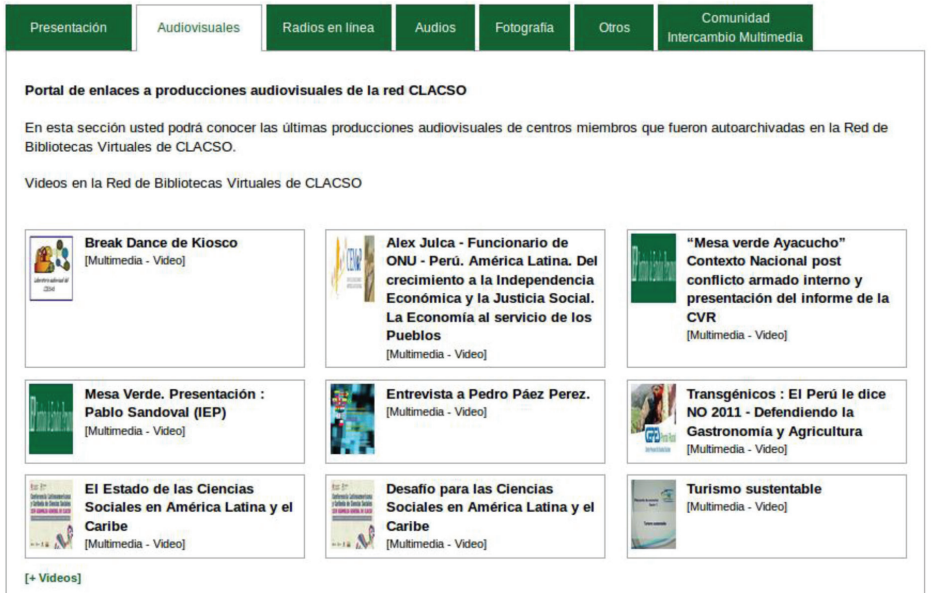
Figure A1
Screen Shot of CLACSO's Reading Room

• **The Multimedia Collection:** The multimedia website (Figure A2) was launched in 2012 and hosts metadata of productions of audio, video, broadcast radio programs, images, etc., of Member Centers, providing

9 http://www.clacso.tv

links to the file hosted on the server of each Center, Vimeo, Youtube, etc. It also provides links to the multimedia web pages of Member Centers.

Figure A2
Screen Shot of CLACSO's Multimedia Collection



• **CLACSO´s Peer-Reviewed Journals Collection**[10]
CLACSO and Redalyc-UAEM have signed an agreement in 2009 so that peer-review journals from CLACSO´s network may apply to Redalyc to be evaluated for inclusion in CLACSO´s peer-review journal collection in Redalyc.

This agreement provides journal editors from CLACSO´s network with Redalyc orientation on best practices for open access publishing and quality editorial processes. Redalyc indexes journals accepted in the evaluation process, and each journal accepted receives from Redalyc indicators on growth, use and profile of scientific production. All indicators prepared by Redalyc-FRACTAL[11]

10 http://clacso.redalyc.org/

11 http://redalycfractal.org/

Figure A3
Screen Shot of CLACSO's Peer-Reviewed Journals Collection

To allow the monitoring of the growth and characteristics of the three digital collections mentioned above, the following indicators displayed in CLACSO´s Digital Repository have been developed:

- Growth of digital collections
- Classification and percentage by type of digital objects
- Topics that describe the content of the Digital Repository of CLACSO
- Indicators of peer-reviewed journals of CLACSO in Redalyc

Each of these indicators is explored in more detail in the following section.

GROWTH OF DIGITAL COLLECTIONS

To follow the growth of digital collections, the following sub-indicators are used:

- Digital objects
- Classification by type and percentage of digital objects
- Authors of digital objects
- Subject descriptors of the collection
- Indicators of CLACSO´s collection of peer-review journals in Redalyc

This data is retrieved from the self-archiving system and the Redalyc reports. The system allows viewing the growth of the general collection at certain periods, as illustrated in Figure A4.

Figure A4
Screen Shot of Data Retrieved on Growth of Digital Collections

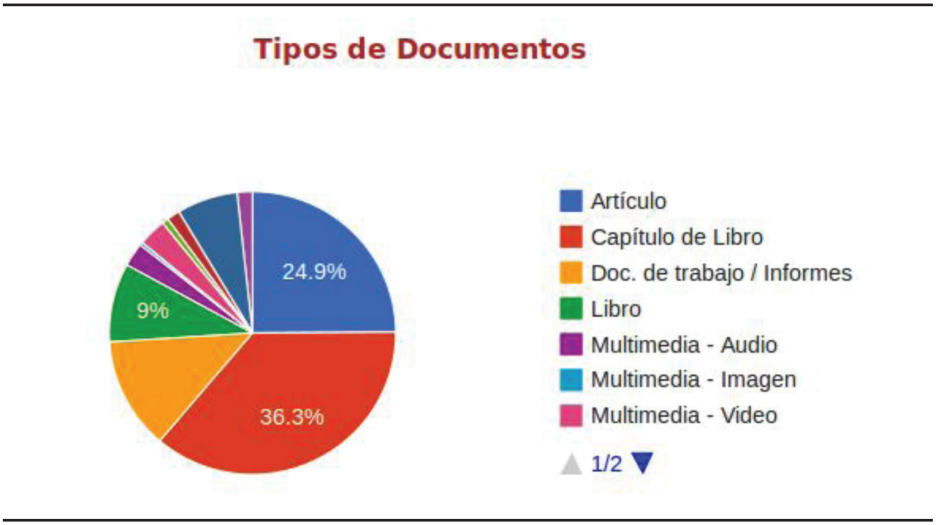
3.1. CLASSIFICATION BY TYPE AND PERCENTAGE OF DIGITAL OBJECTS

The CLACSO website reports the type of digital objects of the entire collection and their percentages. It does not include CLACSO´s collection in Redalyc, which has its own indicators (described in another section). Figure A5 shows the number of documents, authors, and topics that were admitted between January 2013 and December 2013, and corresponds to the digital objects deposited in the Reading Room digital collections and the new system for online self-archiving: documents (1,111), authors (2,435) and subject topics (2,095).

Figure A5
Screen Shot of Data Retrieved on Classification of Digital Objects by Type

Figure A6 shows the percentages of digital objects in CLACSO´s Digital Repository according to type of digital objects deposited: 45.3% corresponds to books and book chapters, 24.9% to journal articles, 12.7% to working documents and research reports. It does not include CLACSO´s peer-review journal collection in Redalyc (described in another section).

Figure A6
Screen Shot of Data Retrieved on Classification of Documents by Type



This graph shows, from the total collection of digital objects in CLACSO´s digital repository, percentages according to type of digital objects deposited: 45,3% corresponds to books and book chapters, 24,9% to journal articles, 12,7% to working documents and research reports. It does not include CLACSO´s peer-review journal collection in Redalyc, which receives indicators described in another section.

3.2. SUBJECT DESCRIPTORS OF THE COLLECTION

From the self-archiving online form, subject descriptors are provided to allow follow-up of subjects from digital objects being incorporated in the Repository. This information is used for having a general mapping of the predominant themes in each period.

Whenever a digital object with its metadata is entered in the self-archiving web-form, and approved, the table with assigned descriptors is updated automatically. Data for the tag-cloud displayed on the Repository website are also taken from that same table.

3.3. INDICATORS OF CLACSO´S COLLECTION IN REDALYC

CLACSO´s network publishes 394 journals, 68% of which are open access through CLACSO´s collection in Redalyc, CLACSO´s Digital

Repository, SciELO and/or within the journals websites. The collection of peer-reviewed journals from CLACSO´s network in Redalyc[12] through December 2013 are as follows:

- 63 journals
- 22,887 articles
- 25,046 authors

Redalyc provides, in its website and/or in special reports, scientometric and bibliometric indicators at the journal, institution of affiliation of the author, and country levels. The Redalyc Fractal Scientometric Lab[13], that prepares scientific activity reports by institution and country, generates these indicators (described in another section).

3.4. INDICATORS OF USE OF CLACSO´S DIGITAL REPOSITORY

Online information is provided on:

- Number of downloads
- Countries that use the service more frequently
- The most downloaded digital objects
- Indicator of use at article level
- Popular subjects

3.5. NUMBER OF DOWNLOADS

The total number of downloads within a given period is available online. This data includes downloaded files from a given path with a particular extension (i.e., .pdf, .doc, etc.).

The statistical system allows differentiating the information of the total number of downloads that were made by users – as a result of searching in CLACSO´s Repository search facility and in search engines – from the traffic generated by robots harvesting the Repository.

In Figure A7, the bottom line corresponds to downloads received from users queries in CLACSO´s search engine and in other search engines. The mid line corresponds to downloads as a result of being harvested regularly. The top line corresponds to the sum of both. Looking at the number of downloads by us-

12 http://clacso.redalyc.org

13 http://redalycfractal.org

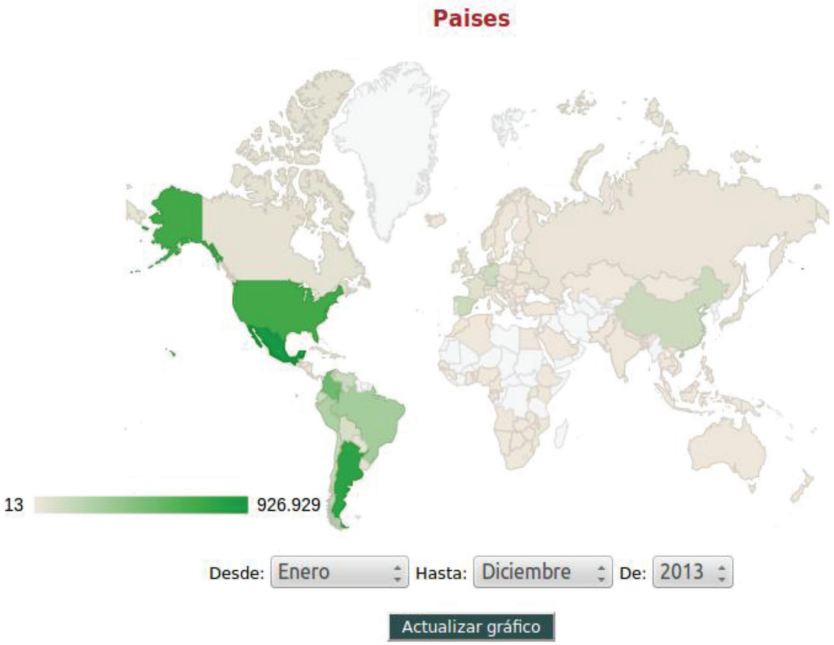
ers, the figure shows that for December 2013 a total of 463.472 downloads were received in CLACSO´s Digital Repository. To this number should be added downloads in CLACSO´s peer-review journal collection in Redalyc, which average 406.000 monthly downloads. So, in total, an estimated average of approximately 850.000 downloads a month.

Figure A7
Screen Shot of Analysis of Number of Downloads

3.6. COUNTRIES THAT USE THE SERVICE MORE FREQUENTLY

Tools allow users to visualize through the map the countries that download digital objects within a certain period. Historically, the countries that more often download digital objects from the Repository belong to Latin America. Mapping downloads by country (Figure A8) indicates that there are other countries outside the Region that download a substantial number of digital objects: United States, Spain, Germany, China, Italy, United Kingdom, France and Japan, among others.

Figure A8
Screen Shot of Analysis of Number of Downloads Mapped by Country



3.7. THE MOST DOWNLOADED DIGITAL OBJECTS

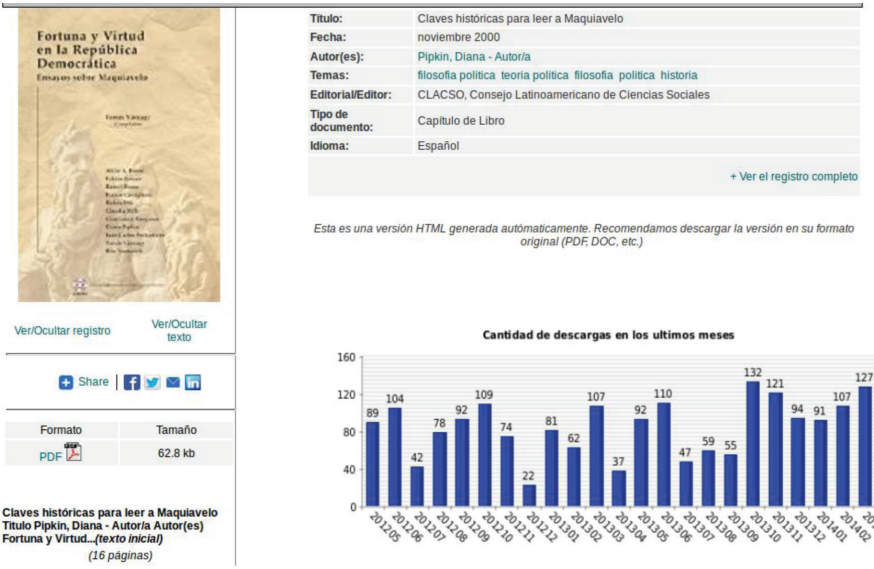
The Repository provides a list of the 30 most downloaded documents in a given month (Figure A9).

3.8. INDICATOR OF USE AT ARTICLE LEVEL

At the end of 2013, CLACSO´s Digital Repository started with the development of indicators at article level to know the use of every digital object of the Repository (Figure A10). This service is in an initial stage of development and testing for a collection of CLACSO-Executive Secretariat digital objects – mainly books and book chapters (3,689 digital objects) – and it will be generalized to the remaining digital collections in the medium term.

Figure A9
Screen Shot of Analysis of Number of Downloads in a Given Month

Figure A10
Screen Shot of Article Level Indicators



3.9. EXAMPLE OF A BOOK CHAPTER WITH INFORMATION ABOUT DOWNLOADS

At the testing stage, it has been found that the individual statistics of some digital objects does not match the numbers obtained with the other indicators of collection. This is because at the time of migration to a new platform a few years ago, it was decided to preserve with the previous URL digital objects for each article, which received many links in their previous URL. Redirection or other strategies are still needed in the system in a next phase. For clarification, this situation occurs only with the digital objects that are hosted in the Repository, not with harvested collections or with CLACSO´s collection in Redalyc.

3.10, POPULAR SUBJECTS

This information comes from the searches performed by users in the Digital Repository search engine, coupled with the searches in search engines (e.g, Google, Yahoo, Bing, etc.). It allows generating tag-clouds of the most requested topics at certain periods. For each request, the web server registers the URL, to determine if the search was undertaken within CLACSO´s search engine or from external search engines. Information is available about the search terms within the URLs.

Figure A11
Screen Shot of Tag-Cloud

This information is complemented with reports from Redalyc about CLACSO´s journal collection in Redalyc.

The development of indicators is a continuing learning process for repositories, mainly for those indicators that are not incorporated as part of the services offered by the platform chosen to manage the Repository.

4. REFERENCES

Alperin, J. P., Fischman, G. E., & Willinsky, J. (2011). Scholarly communication strategies in Latin America's research-intensive universities. *Revista Educación Superior Y Sociedad*, *2*(16). Retrieved from http://ess.iesalc.unesco.org.ve/index.php/ess/article/viewArticle/409 (accessed 8 May 2014)

Archambault, E., & Larivière, V. (2010). The limits of bibliometrics for the analysis of the social sciences and humanities literature. *World Social Science Report 2010: Knowledge Divides,* 251-254. http://unesdoc.unesco.org/images/0018/001883/188333e.pdf#page=303 (accessed 8 May 2014)

Babini, D., y Fraga, J. (comps.) (2006). *Edición electrónica, bibliotecas virtuales y portales para las ciencias sociales en América Latina y el Caribe***.** Buenos Aires: CLACSO. http://biblioteca.clacso.edu.ar/clacso/biblioteca/20110809013130/Babinientero.pdf (accessed 8 May 2014)

Babini, D., González, J., López, F. A., y Medici, F. (2010). Construcción social de Repositorios Institucionales. *Información, Cultura y Sociedad, 23.* http://www.scielo.org.ar/pdf/ics/n23/n23a04.pdf (accessed 8 May 2014)

Babini, D., López, F. A., y Archuby, G. (2011). Indicadores de comunidad, crecimiento y uso para la evaluación de un repositorio institucional. In *Panel Hacia la evaluación de los repositorios institucionales, Segundo Taller de Indicadores de Evaluación de Bibliotecas,* National University of La Plata, Argentina, June 27-28, 2011**.** http://tieb.fahce.unlp.edu.ar/descargables/ppts-presentaciones/babini_lopez_archuby (accessed 8 May 2014)

Babini, D., López, F. A., Medici, F., González, J., y Vergara, F. (2010). Red de Bibliotecas Virtuales de Ciencias Sociales de América Latina y el Caribe, CLACSO. Estrategia frente a las desigualdades regionales. Oportunidades de visibilidad internacional de producción propia. Acceso a la producción de otros países de la región. *Greenstone. Un software libre de código abierto para la construcción de bibliotecas virtuales Experiencias en América Latina y el Caribe.* Montevideo: UNESCO http://

unesdoc.unesco.org/images/0018/001887/188719s.pdf . (accessed 8 May 2014)

Babini, D., Vergara-Rossi, F, y Archuby, G. (2007). Estadísticas de usuarios en una biblioteca virtual. El caso de la biblioteca virtual de la red Clacso. *El Profesional de la Información (Madrid), 16(1),* http://www.elprofesionaldelainformacion.com/contenidos/2007/enero/07.pdf (accessed 8 May 2014)

Babini, D., Vergara-Rossi, F., González, J., y Medici, F. (2008). Biblioteca virtual cooperativa descentralizada con software libre Greenstone. *El Profesional de la Información (Madrid), 17*(1). http://www.elprofesionaldelainformacion.com/contenidos/2008/enero/07.pdf . (accessed 8 May 2014)Cetto, A. M., Alonso-Gamboa, J. O., & Córdoba González, S. (2010). Ibero-American systems for the dissemination of scholarly journals: A contribution to public knowledge worldwide. *Scholarly and Research Communication, 1*(1), 1-16. http://www.src-online.ca/index.php/src/article/viewFile/17/31 . (accessed 8 May 2014) Guédon, J-C. (2008). Open access and the divide between "mainstream" and "peripheral" science. (unpublished). http://eprints.rclis.org/bitstream/10760/10778/1/Brazil-final.pdf . (accessed 8 May 2014)

As regional open access initiatives from Latin America have begun to produce and share indicators, this book on "Open Access Indicators and Scholarly Communications in Latin America" is a first attempt to systematically explore and describe them within the broader context of Open Science. In this process, the publication aims at identifying the presence, growth, use and reach of research results that are now "open" and freely available in the Web.

This book is the result of a joint research and development project supported by UNESCO and undertaken in 2013 by UNESCO (Bhanu Neupane) in partnership with the Public Knowledge Project (PKP); the Scientific Electronic Library Online (SciELO); the Network of Scientic Journals of Latin America, the Caribbean, Spain and Portugal (RedALyC); Africa Journals Online (AJOL); the Latin America Social Sciences School-Brazil (FLACSO-Brazil); and the Latin American Council of Social Sciences (CLACSO) . This book aims to contribute to the understanding of scholarly production, use and reach through measures that are open and inclusive. This publication serves as an important milestone in the UNESCO's Open Access Strategy that 196 countries have collectively endorsed. The publication has also received significant inputs from the PKP-FLACSO-IDRC's project "Quality in Open Scholarly Communication of Latin America", which was undertaken in 2012-2013 in collaboration with Latindex, SciELO and Redalyc.

The present book is divided in two sections. First section presents a narrative summary of Open Access in Latin America, including a description of the major regional initiatives that are collecting and systematizing data related to Open Access scholarship in the region, and of available data that can be used to understand the (i) growth, (ii) reach, and (iii) impact of Open Access in developing regions. The first section ends with recommendations for future activities. The second section includes in-depth case-studies with the descriptions of indicators and methodologies of peer-review journal portals SciELO and Redalyc, and a case of subject digital repository maintained by CLACSO.

Sida
Swedish International
Development Cooperation Agency

Consejo Latinoamericano
de Ciencias Sociales
CLACSO
Conselho Latino-americano
de Ciências Sociais
Latin American Council of Social Sciences

9 789877 220278