

Article

Scientific Production on Open Access: A Worldwide Bibliometric Analysis in the Academic and Scientific Context

Sandra Miguel ^{1,†}, Ely Francina Tannuri de Oliveira ^{2,†} and Maria Cláudia Cabrini Grácio ^{2,*}

Received: 15 October 2015; Accepted: 5 January 2016; Published: 14 January 2016

Academic Editor: Beatriz Valadares Cendón

¹ Instituto de Investigaciones en Humanidades y Ciencias Sociales—IdIHCS (UNLP-CONICET), m, Universidad Nacional de La Plata, calle 51 e/124 y 125, Ensenada, 1925, Argentina; smiguel@fahce.unlp.edu.ar

² UNESP—Univ Estadual Paulista, 737 Hygino Muzzi Filho Avenue, 17525-900 Marília, Brazil; etannuri@gmail.com

* Correspondence: cabrini@marilia.unesp.br; Tel.: +55-014-99810-4566; Fax: +55-14-3402-1302

† These authors contributed equally to this work.

Abstract: This research aims to diachronically analyze the worldwide scientific production on open access, in the academic and scientific context, in order to contribute to knowledge and visualization of its main actors. As a method, bibliographical, descriptive and analytical research was used, with the contribution of bibliometric studies, especially the production indicators, scientific collaboration and indicators of thematic co-occurrence. The Scopus database was used as a source to retrieve the articles on the subject, with a resulting corpus of 1179 articles. Using Bibexcel software, frequency tables were constructed for the variables, and Pajek software was used to visualize the collaboration network and VoSViewer for the construction of the keywords' network. As for the results, the most productive researchers come from countries such as the United States, Canada, France and Spain. Journals with higher impact in the academic community have disseminated the new constructed knowledge. A collaborative network with a few subnets where co-authors are from different countries has been observed. As conclusions, this study allows identifying the themes of debates that mark the development of open access at the international level, and it is possible to state that open access is one of the new emerging and frontier fields of library and information science.

Keywords: open access; scientific production on open access; bibliometric analysis

1. Introduction

For centuries, the model of science communication was based on the publication of articles in journals, first printed and then also electronically, where a subscription fee for access to the articles was the predominant system of distribution and circulation of publications. However, over the past decade, this model has been going through profound changes, especially motivated by the emergence of open access (OA) to scientific knowledge.

OA emerged earlier this century, partly as a result of the phenomenon known as the “crisis in scientific publishing”, which occurred as a consequence of high prices for subscriptions, libraries' reduced budgets and consequent cancellation of subscriptions and restrictions on access to the articles by the scientific community [1,2]. Initially formalized through the Budapest Declaration in December 2001 [3] and followed a few years later by Berlin [4] and Bethesda [5], OA proposes free access to scientific publications, as an alternative to the traditional subscription model to journals and payment for access to the articles. One of the basic principles of OA is to allow users to read, download, copy,

distribute, print, search or link to full texts of articles, without economic, legal or technical barriers other than the Internet itself. The only condition established is to give authors control over the integrity of their work and the right to be properly acknowledged and cited.

In a recent study, Björk *et al.* [6] summarized the main arguments supporting the principles of OA. One is to increase the visibility of research results contributing to faster advancement of science. Another is that since much of science is funded with public funds, its results should be publicly accessible. In addition, one of the expected consequences in the scientific publication system is an overall reduction in the costs of publication and dissemination processes, compared to the subscription model.

From the beginning, two venues to deliver open access have been established: the golden road, which proposes the publication of articles in open access journals, and the green road, which proposes the self-archiving of published articles in traditional journals of subscription in open access digital repositories. Graziotin, Wang and Abrahamsson [7] present the concepts and definitions involved in OA in detail.

In 2012, the Budapest Open Access Initiative [3] reaffirmed the principles of OA, as well as the two developing venues, highlighting that not only they are direct and effective means of achieving open access, but also they are available to the academic community immediately and avoid publishing markets or legislation delays. However, major publishers worldwide have adopted OA models that do not ensure immediate open access to articles in all cases, imposing deadlines for availability in repositories or in other cases transferring the subscription costs to access through payment by the author (hybrid journals) or payment for article processing costs (APC -article processing charges).

Although the future of OA is uncertain, some authors, such as Harnad *et al.* [8], suggest that the green road is the only alternative to reach 100% OA. However, “what, when and where” self-archiving takes place is controversial [9]. Other authors argue that golden OA has fared better and has more potential for economic stability than green OA [10]. Suber [11] suggests that both roads of OA should complement each other. The truth is that the volume of articles available in open access is still low. In an overview of recent studies, Björk *et al.* [6] reported that the percentage of open access articles in journals included in the Web of Science (WoS) and Scopus is barely 23% (on the two roads), although their level of development differs in thematic areas and geographical regions [12–14].

The importance this issue has gained worldwide in different sectors related to the development of science and scientific communication, involving authors, editors, librarians, scientific policy managers, among others, has led to a proliferation of studies that make up a large volume of scientific production regarding OA.

However, there are few papers that conduct a literature analysis from a bibliometric perspective, analyzing international publications and with a multidisciplinary coverage. One article constitutes an investigation on the structure and evolution of the theme of open access in journals included in the WoS from 2000 to 2010, based on co-citation analysis. The study finds that open access is an emerging field. The first three prominent papers are Lawrence’s “Free online availability substantially increases a paper’s impact” (2001), Antelman’s “Do open access articles have a greater research impact” (2004) and Swan and Brown’s “Open access self-archiving: An author study” (2005). Harnad is the most influential author, and *Journal of the Association for Information Science and Technology* (JASIS/JASIST) is the most frequently cited journal in the open access literature [15]. Other studies carry out a literature review examining the relationship between OA status and citation counts of scholarly articles [16].

Considering that the subject OA was more disseminated from 2002, there are few studies that focus worldwide on the directions and trends of the theme, analyzing diachronically its presence in mainstream science, identifying the countries that have taken the lead in the development of research on this theme, as well as the elite researchers in the world and core science journals in the dissemination of this knowledge. Current bibliometric studies on this topic have been dedicated to regional analysis, by area or discipline, thus failing to provide a worldwide state-of-the-art visualization on the theme.

In this context, this research aims to diachronically analyze the worldwide scientific production on open access, in the academic and scientific context, in order to contribute to knowledge and visualization of its main actors.

More specifically, this study proposes to:

- Diachronically analyze the presence of the theme open access in the international scientific literature;
- Identify the leading countries and main authors;
- Explicate the core journals disseminating the theme;
- View the scientific collaboration network among the most productive authors;
- Identify the main sub-themes pertaining to the theme and their co-occurrences in the analyzed literature, through the words present in titles and abstracts of articles.

2. Data and Methods

This study is a bibliographical, descriptive and analytical research, with the contribution of bibliometric studies, especially the production indicators, scientific collaboration and indicators of thematic co-occurrence.

Traditionally, the source of information used for bibliometric studies at a worldwide level was the joint set of databases belonging to the Web of Science (WoS) of Thomson Reuters. One of their strong points is their multidisciplinary and international nature. These databases contain information from some 12,000 scientific journals and have become tools of world reference for information retrieval and for studies evaluating science. However, the appearance on the market of the Scopus database, by Elsevier, with coverage of over 20,000 journals, has helped complement and broaden analyses drawn from other sources. Ever since its early days, the Scopus database has been a subject of analysis [17–19], with comparisons of coverage, accessibility, usability and price between this tool and the WoS [20,21] or comparing these two products with Google Scholar [22,23]. We shall avoid exhaustive mention of the profuse literature surrounding Scopus and simply note that certain limitations [21] accompany the tremendous endeavor undertaken by Elsevier. There are still omissions of journals regarding coverage and the affiliation country. This notwithstanding, the levels of chronological, geographic and thematic coverage attained by Scopus are indeed sufficient for guaranteeing relevant and reliable findings [13,24,25].

Using Scopus database as source, on 2 September 2015, the retrieval of articles on the subject was carried out, using as a strategy the following terms in Advanced search, denominated **Strategy**:

(TITLE (“open access” OR “open archives” OR “free access”) OR KEY (“open access” OR “open archives” OR “free access”)) AND TITLE-ABS-KEY (scholarly OR publishing OR scientific OR libraries OR repositories OR academic) AND DOCTYPE(ar) AND PUBYEAR < 2015 AND PUBYEAR >1900 where:

TITLE: search in the field title

KEY: search in the field keyword

TITLE-ABS-KEY: search in the fields title, abstract and keyword simultaneously

DOCTYPE(ar): limits the search to the type of document “Articles”

PUBYEAR < AND PUBYEAR >: limits the search to a period of time according to the year of publication

Thus, these terms allowed retrieving articles in which one of the terms “open access” or “open archives” or “free access” appears in the title or among the keywords. To contextualize these terms, in the context of scientific publication, the search was restricted for the presence of the terms scholarly or publishing or scientific or libraries or repositories or academic, for the title, abstract or keyword. We highlight the nearness of this search strategy in comparison to the one proposed by the PICo method,

a methodology used in non-clinical research, in which P means population, I means interest and Co is for context [26]. In this study, P corresponds to the articles on the thematic OA, represented in the search strategy by the terms TITLE and KEY. Thus, I represents the interest in knowing the diachronic evolution of OA in mainstream science. Finally, Co is represented by the terms TITLE-ABS-KEY, DOCTYPE(ar) and PUBYEAR. The search was filtered excluding articles published in 2015 and those that had the term "free access", but were not on the subject. The resulting corpus was of 1179 articles. In all of the other retrieved articles, the studied theme was observed. This set was exported in Research Information Systems (RIS) format, with the selection of the following fields: author(s), title, year, source title, affiliation, abstract and keywords.

Next, using Bibexcel software, frequency tables for the variables authors, year and journal were constructed. For the diachronic analysis of the theme worldwide, the frequency table for the year of publication was converted into a line chart, through Excel software.

For the construction of Table 1, related to the authors' countries of origin, as well as the analysis of the most recurring keywords, the Scopus report was used.

Table 1. Major producing countries on the theme open access.

Country	Article Number	%
United States	358	30.4
United Kingdom	158	13.4
Spain	70	5.9
Germany	67	5.7
Canada	58	4.9
India	54	4.6
Australia	38	3.2
France	33	2.8
The Netherlands	29	2.5
Italy	26	2.2
Brazil	26	2.2
China	21	1.8
Finland	21	1.8
Austria	20	1.7
Switzerland	20	1.7
Belgium	17	1.4
Greece	17	1.4
Nigeria	15	1.3
Iran	14	1.2
Denmark	14	1.2
South Africa	14	1.2
Sweden	13	1.1

To visualize the recognition of the most productive authors (Table 2) from the examined community, the total number of articles was analyzed in which each author was cited, among the retrieved ones. For each author, this citation retrieval was carried out adding the search term REFAUTH (Author name) in **Strategy**, in the advanced search.

Table 2. Most productive researchers on the theme open access.

Researcher (Country)	Article Number
Björk, Bo-Christer (Finland)	18
Xia, Jingfeng (United States)	10
Davis, Philip M. (United States)	8
Harnad, Stevan (Canada)	8
Schopf, Joachim (France)	7
Solomon, David J. (United States)	7
Abadal, Ernest (Spain)	6
Bauer, Bruno (Austria)	6
Laakso, Mikael (Finland)	6
Melero, Remedios (Spain)	6
Moskovkin, Vladimir M. (Russian Federation)	6
Mukherjee, Bhaskar M. (India)	6
Abrizah, Abdullah (Malaysia)	5
Beall, Jeffrey (United States)	5
Jacsó, Péter (United States)	5
Kousha, Kayvan (Iran)	5
McGrath, Mike (United Kingdom)	5
Nicholas, David (United Kingdom)	5
Oppenheim, Charles B. (United Kingdom)	5
Rodrigues, Rosângela S. (Brazil)	5
Sánchez Tarragó, Nancy (Cuba)	5
Willinsky, John (United States)	5

Regarding the journals in Table 3, the SCImagoJR Portal was used for the record country of issue, as well as the quartile to which each belongs. To verify the OA form of these journals, the Portal SHERPA/RoMEO-Rights Metadata for Open archiving was used, which allows visualization of the OA road type by journal or by editor, categorizing the archiving policy into: green (can archive pre-print and post-print or publisher's version/PDF), blue (can archive post-print (*i.e.*, the final draft post-refereeing) or publisher's version/PDF), yellow (can archive pre-print (*i.e.*, pre-refereeing) and white (archiving not formally supported).

For the construction of the co-authorship network of the most productive researchers, Scopus was used, in order to retrieve their co-authors. For this search, the term AUTH (Author name) was added to Strategy, in advanced search.

For the set of coauthors retrieved for each most productive researcher in Table 2, the co-authorship network of coauthors with at least two articles published with the researcher was generated. This criterion was selected due to the fact that there are 101 co-authors responsible for just one article among the total, which would hinder the visualization of the consolidated collaboration on this theme. Pajek software was used for the construction of the network.

VoSViewer software was used to build and visualize the co-occurrence network of keywords in order to evaluate the co-occurrence of these words in titles and abstracts. In total, 284 different terms were extracted, with which the co-occurrence network was built. It was decided to leave the keywords as recorded by the authors to see the differences in the vocabulary, in order to identify changes and predominance. Maps were created using the VoS mapping technique and the VoS clustering technique, with colors distinguishing the clusters.

Table 3. Core journals on the dissemination of the analyzed scientific production.

Journal (Country of Editor)	Article Number	%	Quartile	Via
Serials Librarian (U.K.)	34	2.9	Q2 Q1	Green
Learned Publishing (U.K.)	23	2.0	Q1	Blue
J. Am. Soc. Inf. Sci. Technol (USA)	23	2.0	Q1	Green
Information Services and Use (Netherlands)	22	1.9	Q3	Green
Serials Review (U.K.)	21	1.8	Q2	Green
El Profesional de la Información (Spain)	20	1.7	Q2	Blue
Insights (U.K.)	20	1.7	Q3	— ¹
OCLC Systems & Services: International digital library perspectives (U.K.)	19	1.6	Q3	Green
Journal of Academic Librarianship (Netherlands)	18	1.5	Q1	Green
First Monday (USA)	18	1.5	Q2	Blue
VOEB-Mitteilungen (Austria)	16	1.4	Q4	— ¹
Online Information Review (U.K.)	16	1.4	Q2	Green
D-Lib Magazine (USA)	16	1.4	Q1	Blue
TripleC (Austria)	14	1.2	Q2	Green
Scientometrics (Netherlands)	14	1.2	Q1	Green
Interlending and Document Supply (U.K.)	14	1.2	Q2	Green
LIBER Quarterly (Netherlands)	13	1.1	Q3	Green
Information Research (U.K.)	12	1.0	Q3	Blue

¹ Journal not found in the Portal SHERPA/RoMEO.

3. Results and Discussion

3.1. Evolution of the Theme Open Access in the International Scientific Literature

Figure 1 shows the trend in the evolution of the theme worldwide from the beginning of the discussions until 2014, when an upward curve is observed, with some years corresponding to inflection points. By 2012, the scientific literature on the subject shows a small fall.

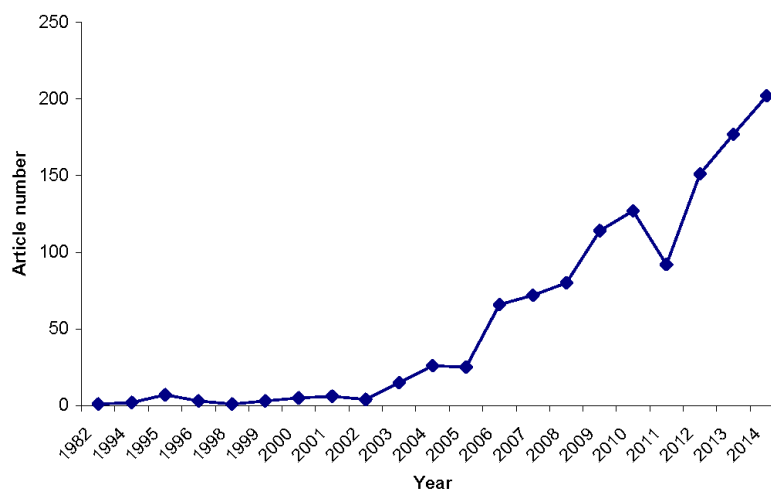


Figure 1. Evolution of the theme "open access" in Scopus.

As mentioned in the Introduction, in December 2001, the Open Society Institute conference took place in Budapest, when the term open access was coined and formalized. Then, in January 2002, the declaration of the Budapest Open Access Initiative (BOAI) set out the principles, strategies, rules and commitments related to open access to research literature.

These facts explain the first point of inflection of the curve, from 2002, gaining strength in 2006 with the launch of PLoS One. In 2008, the purchase of the OA publisher BioMed Central, by Springer, one of the market leaders in academic journal publishing, consolidated the OA model [27].

Since then, publications on this subject show a sharp rising curve, with the exception of 2011, when it presents a significant decrease, which deserves a more detailed study in order to identify the causes.

3.2. Countries and Main Authors of the Theme Open Access in the International Scientific Literature

Table 1 shows the 22 major producing countries from a total of 74 countries that have published on the topic open access. Each country produced at least 13 articles in the period, corresponding to at least 1% of the total production of articles on the theme and their respective percentages.

It is observed that the countries presented in Table 1, in general, are the most prominent in world science, with some exceptions, such as Japan, which, according to the SCImagoJR Portal, ranks fifth in the world rankings and is not found among the main producing countries on open access, as well as South Korea, the Russian Federation and Taiwan. On the other hand, Spain, The Netherlands and Brazil occupy top rankings in the scientific literature on open access in relation to those occupied in world science, suggesting that these countries, especially Spain, have undertaken a major effort in the development of the analyzed subject.

Regarding the directions of open access strategies, it is emphasized that the United States has privileged the green road. In 2013, the Fair Access to Science and Technology Research (FASTR) Act demanded open access to articles resulting from major federal agencies funding via repositories. In their scientific production, the most recurrent keywords indexing the articles are: articles (62), access to information (48), bibliometrics and keywords related to its procedures (31), journals (30), database (30) and repositories (22), among others. Given the emphasis on the keyword “articles”, we point out the alignment between the produced scientific literature and the privileged open access strategy adopted by the country (green road).

The United Kingdom, through the Research Council, favors open access via the golden road, by largely incentivizing publishers. In its scientific production, a wide dispersion of indexing keywords of articles is observed, highlighting the following: articles (31), access to information (23) and journal (15), among others.

In Spain, Law 14/2011 on Science, Technology and Innovation (Boletín Oficial del Estado, 2 June 2011), in its Article 37 that addresses the dissemination in open access, makes it clear that public officials of the Spanish System of Science, Technology and Innovation must encourage the development of open access repositories for publications belonging to their researchers and must establish systems that allow connecting them to similar initiatives nationally and internationally. Its scientific production emphasizes the following keywords in the articles: repository (21), access to information (15), journals (22), articles (14) and bibliometrics and keywords related to its procedures (9).

In The Netherlands, twenty institutions, including the KNAW—Royal Netherlands Academy of Arts and Sciences, VSNU—Association of Universities in the Netherlands and The Netherlands Association of Universities of Applied Sciences (Vereniging Hogescholen) signed the Berlin Declaration [28]. The Dutch government intends to change completely to open access via the golden road, within approximately ten years, *i.e.*, by 2024. In line with this proposal, the NOW—Netherlands Organization for Scientific Research, the main funder of scientific research in Dutch universities and institutions, primarily adopts the golden OA program for research funds and, where this is not possible, supports the green OA program [29]. The following keywords stand out in the scientific

production of this country: journals (5), article (4) and repository (2), with a large dispersion of terms indexing the articles.

In Brazil, Bill 387/2011 is based on the green model. Through this project, higher education institutions and research units must build institutional open access repositories with the use of internationally-standardized metadata. The following keywords indexing the articles are: journals (6), access to information (4) repository (3) and articles (2).

In relation to the authors responsible for the 1179 articles, we found a total of 2407 authors, which results in an average of 2.04 authors per article. It is noteworthy that 2108 authors have produced one single article (87.6%), suggesting a large dispersion of researchers producing scientific literature on the theme.

Table 2 presents the 22 most productive researchers, in descending order of quantity of published articles, with the publication of at least five articles on the topic. This criterion was adopted, since 50 authors published four articles, indicating a large dispersion.

It is initially highlighted that the 22 researchers presented in Table 2 were responsible, together, for 132 (11.2%) articles of the total of the analyzed articles, all of which are published from 2003. Thus, the 34 articles that comprise the period 1982–2002 were from less quantitatively-expressive authors, considering that they published, individually, fewer than four articles.

Analyzing Table 2, researchers from different countries are found, but especially the United States, the United Kingdom, Spain and Finland. Among the researchers with the highest number of articles, pioneer and precursor authors on the theme open access are found, such as Stevan Harnad, who is found in the study of Duzyol *et al.* [15].

Among the 22 researchers presented in Table 2, 13 have significant influences on the set of 1179 analyzed articles, being cited in at least 1% of these articles, namely: Harnad (20.4%), Oppenheim (12.4%), Björk (10.6%), Davis (8.3%), Willinsky (7.8%), Laakso (6.6%), Nicholas (5.9%), Xia (3.6%), Solomon (3.6%), Kousha (3.1%), Beall (2%), Jacsó (1.6%) and Melero (1.4%).

Stevan Harnad, defender and activist for the green road, from Université du Québec à Montréal, was the advocate of open access in the 1990s by proposing that all scientific literature were to be available on the Internet for free and unrestricted access and proposing the names of the strategies green road, golden road and self-archiving [30]. His articles, retrieved in this research, were published in the 2005–2013 period, individually or in co-authorship with researchers from Canada, the United Kingdom and France. He addresses, especially, issues related to cognitive science and open access, which have contributed with the grounds of the research of many investigators from different countries from all continents. Within the analyzed literature, he was the most cited author (in 240 of 1179 analyzed articles) in the scientific production of researchers from 41 different countries, especially the USA, the U.K., Spain, India, Canada, Finland, Germany, Iran and Brazil, produced in seven different languages, especially English, Spanish and Portuguese. This result is aligned to what Duzyol, Taskin and Tonta [15] found in identifying Harnad as the most influential author on the subject.

Charles B. Oppenheim, from Cass Business School, London, expert in copyright, data protection, freedom of information, open access and other aspects of the dissemination of research, bibliometrics and research evaluation, with five articles published between 2003 and 2010, is the second most cited author in the analyzed literature (cited in 146 articles) to be referenced by researchers from 37 different countries, especially the USA, the U.K., Spain, India, Canada, Argentina, Japan and Finland, in articles published in six different languages, mainly English and Spanish.

Bo-Christer Björk, researcher at the Department of Management and Organization, Hanken School of Economics, Helsinki, and proponent of the scientific communication model named the global distributed information system (2007), was the author with the highest number of published articles, with almost twice as many articles compared to the second most productive author. His publications, constituting this study, are distributed regularly from 2004–2014, published individually authored or an elite at the front of scientific research in the studied subject.

Philip M. Davis, researcher at Phil Davis Consulting, Ithaca, published eight articles presented in Table 2, individually or co-authored with researchers from his own country more recently (2008–2013), investigating in particular the theme open access with bibliometric analyses, especially citation analysis. He was theoretical referenced to 98 articles, published by authors from different countries (28), especially the United States, the United Kingdom, Spain and Canada.

John Willinsky, researcher from Stanford University, Graduate School of Education, Palo Alto, published five articles individually or co-authored with researchers from his own country and Canada in the 2003–2013 period, addressing especially open access in the medical field. His articles compose the theoretical reference of 92 of 1179 analyzed articles, published by researchers from 27 different countries, mainly the United States, the United Kingdom, Brazil and Spain, published in five different languages, mainly English, Portuguese and Spanish.

Mikael Laakso, from the Department of Management and Organisation, Hanken, Svenska Handelshögskolan, Helsinki, published his articles recently (2010–2014), with six of them in co-authorship with researchers from five different countries, as well as researchers from his country. His articles discuss, especially, meta-theoretical studies on open access, as well as issues related to the green road strategy. His articles form the theoretical reference of 78 of the 1179 analyzed articles published by researchers from 27 different countries, mainly the United States, the United Kingdom, Spain, Canada, Finland and India, published in seven different languages, especially English and Spanish.

David Nicholas, from the Centre for Information Behaviour and the Evaluation of Research/School of Library, Archive and Information Studies at University College London, published all of his articles in co-authorship with researchers from the United States and Iran, as well as his own country's researchers in the area of social and computer sciences in the period 2005–2009. These articles form the theoretical reference of 70 of the 1179 analyzed articles published by researchers from 24 different countries, especially, the United States, the United Kingdom and Australia, published in four different languages, especially English and Spanish.

Jingfeng Xia, from the Department of Library & Information Science, Indiana University (USA), showed significant scientific production on the theme, disseminated from 2009–2012, in particular related to repositories and analysis of the impact of open access scientific production via citation of scientific production. He developed research in individual authorship and in co-authorship with researchers from his country. Among the 1179 analyzed articles, he was cited in 43, especially in articles in the fields of social sciences and computer science, published in English, especially by authors from the United States, Spain, Canada and Germany.

David J. Solomon, from the College of Human Medicine, Michigan State University, published his articles in individual authorship or in collaboration with the Finnish researchers Björk and Laakso, addressing, especially, issues related to open access journals in the medical field and meta-theoretical studies based on bibliometric analyses, from 2006–2014. He was cited in 42 articles, especially by researchers from the United States, Canada, Finland, Spain and the United Kingdom, as well as 26 other countries, mainly published in English and Spanish.

Kayvan Kousha, from the Department of Library and Information Science, University of Tehran, published his articles in individual authorship and also in co-authorship with researchers from the United Kingdom (Mike Thelwall) and Spain from 2007–2014, in the area of social sciences and computer sciences, especially addressing the issue of the impact and citation of OA publications. This scientific production grounded 37 articles of the set of analyzed articles, produced by researchers from seven countries, especially the United States, Canada, India and Iran.

Jeffrey Beall, from the University of Colorado, Denver, individually published articles retrieved in this study, from 2009–2013, in which he presents reflections, debates and controversies around open access, among them the issue of predatory publishers and OA funding via the golden road. This scientific production was referenced in 23 articles, published by researchers from 10 countries, such as the United States, the United Kingdom and Finland.

Péter Jacsó, from the Department of Computer Science and Information System, University of Hawaii, Hawaii, published articles in the period 2003–2010, in individual authorship, discussing technical and methodological aspects of indexing and scientific information retrieval and bibliometric analysis on the theme OA. This scientific production was referenced in 19 articles, published by authors from 13 countries, in particular India, the United Kingdom and the United States.

Remédios Melero, Spanish researcher at the Instituto de Agroquímica y Tecnología de Alimentos (IATA), Valencia, published her articles in the period 2009–2014, in individual authorship and also in co-authorship with other Spanish researchers in the fields of social sciences and computer science, especially in Spanish. Her scientific production grounded 17 articles of the analyzed corpus of researchers from Spain, Argentina and Brazil, Mexico and Malaysia, published in English and Spanish.

It is noteworthy that the nine investigators presented in Table 2 with less impact (cited by less than 1% of the articles) in the analyzed literature, in general, published their articles more recently, from 2009, with most of them published in 2014, the last year of the analyzed period, namely: J. Schopf (2005–2014); E. Abadal (2009–2014); B. Bauer (2009–2014); V. Moskovkin (2008–2014); B. Mukherjee (2007–2014); A. Abrizah (2009–2013); M. McGrath (2012–2013); R. Rodrigues (2010–2014); N. Sánchez-Tarragó (2007–2012). Thus, one of the possible hypotheses for the low impact is the fact that these publications are less likely to have been accessed by the scientific community, while among the authors with the greatest impact, in general, their articles had been disseminated in previous years.

Another possible hypothesis for the low impact relates to the fact that these researchers are from non-English speaking countries and have published individually or, when co-authored, only with collaborators from their own countries or neighboring countries, also non-English speaking countries, such as Germany (Bauer), Serbia and Russia (Moskovkin), Brazil (Abadal), Malaysia and Iran (Abrizah) and Spain (Rodrigues and Sánchez-Tarragó). This characteristic aligns with the observation of several bibliometric scholars who claim that there is a tendency: the higher the international scientific collaboration, the higher the impact of the published literature [31,32].

3.3. Core Journals of the Theme Open Access in the International Scientific Literature

Table 3 shows the core journals on the studied subject, corresponding to 18 titles of the 488 journals that disseminated the retrieved articles. These journals published each at least 1% of the 1179 articles, totaling 333 articles, corresponding to 28.2% of publications.

It is noteworthy that this core of 18 journals is published by a group of five countries; eight of them are of British origin (U.K.). The Netherlands, a country of large commercial publishers who disseminate academic publications, is the country of origin of four journals, all indexed in the field of library and information science. In addition, the United States, the largest producer of articles on the theme (Table 1), is the country of origin of three journals.

Regarding quartiles, 67% of the journals belong to Q1 and Q2 in social sciences, suggesting that journals with higher impact in the academic community have disseminated the new knowledge constructed on the theme.

The majority (11) of the journals are categorized in Portal SHERPA/Romeo under the green archiving policy, *i.e.*, allowing an archive pre-print and post-print or the publisher's version/PDF, and a significant part (five) of them under the blue archiving policy, which allows archive post-print (*i.e.*, final draft post-refereeing) or publisher's version/PDF. The two journals not located on this portal are golden road journals. It is interesting to note that the journals publishing under OA are not mostly in the golden road. A recent study conducted by Ennas and Di Guardo [33] found that the golden road Scopus journals, which are SCImago Journal & Country Rank (SJR) top rank, tend to require a fee to publish; only publishers located in the USA and the U.K. are positively related to the ranking; and only journals publishing in English are positively related to high rankings.

Thus, it is observed that there is a concern from the authors in reconciling the theme under study and the type of journal chosen for the socialization of new knowledge generated in this subject.

3.4. Scientific Collaboration Network among the Most Productive Authors

The scientific collaborative network of the most productive researchers on the theme, together with the co-authors with whom researchers have published more than one article in co-authorship, is presented in Figure 2. The 22 authors highlighted in red correspond to the most productive ones on the theme, and the thickness of segments represents the frequency of co-authorship between two authors.

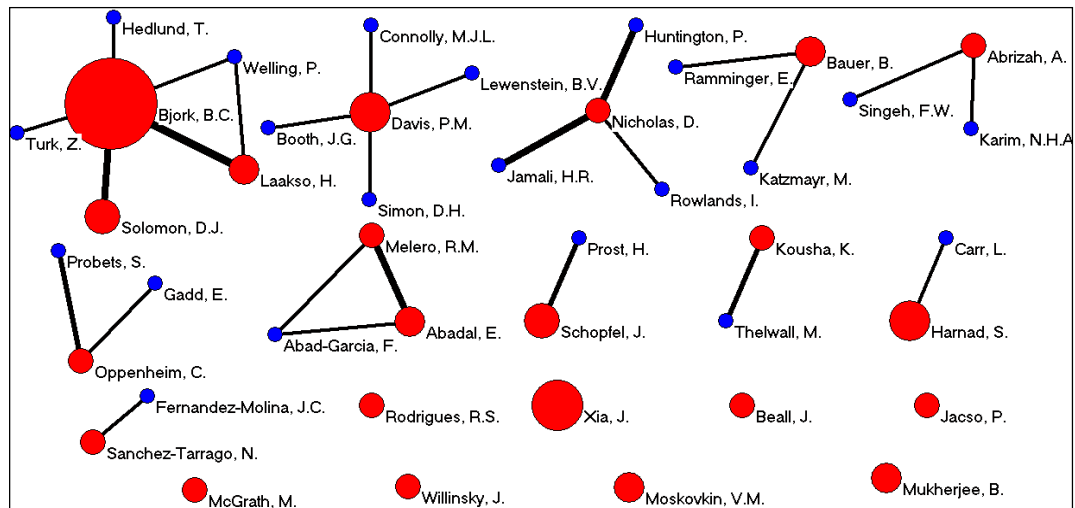


Figure 2. Collaborative network among the 22 most productive authors who had at least two co-authorships.

The network is fragmented into small subnets and does not constitute an organic whole. The author Björk, B-C (Finland) is highlighted in the biggest subnet, as the most productive one, who has co-authorship with five researchers, two of them belonging to the group of most productive authors, namely: Laakso, P. (Finland), with whom he has five co-authorships, the biggest number of co-authorships, and Solomon, D.J. (United States), with whom he has four co-authorships.

It is also noted that author Björk, B-C presents the greatest degree centrality equal to five, as he has the largest number of connections, *i.e.*, with Welling P. (Finland), Hedlund T. (Finland) and Turk, Z. (Slovenia), together with the two most productive authors that were previously mentioned. Considering the geographic distances among the origin of these researchers, it is worth highlighting that author Björk, B-C has wide-ranging relationships, and this subnet can be an indicator of a more dense net creation.

A second subnet highlighted is the triad consisting of the two most productive researchers, Melero, R.M. (Spain) and Abadal, E. (Spain), together with the collaborator Abad-Garcia, F. (Spain). Spain primarily adopts the gold OA.

Although collaborations between Spanish and Brazilian researchers do not appear in the network, it is interesting to highlight a recent publication, of April 2014, by Rosangela Rodrigues and Ernest Abadal, who outline an overview of the open access publishing model in Brazil and Spain. The authors conduct an analysis of the journals from these countries that are indexed by the Web of Science (WoS) and Scopus, according to some aspects of publication format, either online or printed, the type of access, either open and by subscription, and the technology platform used.

Five other small subnets stand out: with Davis, P.M. (United States) in the center of the subnet and the co-authors Conolly, D. (United States), Booth, J.G. (United States), Simon, D.H. (United States) and Lewenstein, B.V. (United States), a typically American subnet, whose country has emphasized the green road.

Nicholas, D. (United Kingdom) has four co-authorships with Jamali, H.R. (Iran) and four with Huntington, P. (United Kingdom), defined by the strongest segments, and with Rowlands,

H. (United Kingdom). It is worth highlighting the homogeneity of this group, with a single researcher coming from an Asian country.

Small subnets, centered on Bauer, B. (Austria) and his two colleagues, Kaminger, E. (Austria) and Katzmayr, M. (Austria), another subnet centered on Abrizah, A. (Malaysia) and his two collaborators, Singeh, F.W. (Malaysia) and Karim, N.H.A. (Malaysia), and, lastly, Oppenheim, C. (United Kingdom) and his two colleagues Gadds, E. (United Kingdom) and Proberts, S. (United Kingdom), are presented. Therefore, there are two subnets with researchers belonging each to the same country.

There are still four dyads that in every case included the most productive authors in co-authorship with other researcher; eight of the most productive authors are isolated. It is noticed that among those, two of them, Jacsó and Beall, work individually; and McGrath, Moskovkin, Rodrigues, Mukherjee, Willinsky and Xia had co-authorships, but with authors of one single work. Therefore, these last authors are not contemplated in the subnet. As conclusions of the network analysis, it is noticed that there are a few subnets where co-authors come from different countries, such as the first and largest subnet centered on Björk, B.C., and the small subnet centered on Nicholas, D. The others are small subnets, consisting of co-authors coming from the same country, showing a very incipient collaborative net, with subnets that scarcely extrapolate to the limits of their own countries, justified by the recent expansion of the matter and the underdevelopment of policies encouraging open access.

3.5. Main Sub-Themes Pertaining to the OA Subject

During the analyzed period, the following subthemes are highlighted from the keywords registered: dissemination and information access (221), articles (192), journals (155), repositories (151), libraries (122), bibliometrics (74)—mostly citation analysis, impact factor and scientific collaboration analysis- and database (66). The publications addressing the strategy named “green road”, in which open access is provided through the self-archiving of scientific researches in repositories, by the author, can be represented by the terms articles, repositories and library. On the other hand, researches related to the strategy named “golden road”, referring to the scientific journals of open access, are represented by the keywords journals and impact factor, among others, associated with the analyzed articles.

The map of the co-occurrence of words in the titles and abstracts of articles presented in Figure 3 shows the main core themes representing the international literature about open access in the analyzed period. It is possible to visualize five clusters distinguished by colors. In each cluster, each node is tagged with a word or sentence extracted from titles and abstracts. Its size is an indicator of its relative importance in the set of nets, and the proximity with other nodes is an indicator of the similarity based on the frequency of the joint appearance of words or sentences in the same title or abstract. The lines connecting the nodes represent the relationship between them.

Cluster 1 is identified with nodes in red, group studies that approach the theme open access from a perspective of changes in the model of scientific communication. Cluster 2 (represented in yellow) includes studies about the attitudes and perceptions regarding OA by the academic community. Cluster 3 (green nodes) represents the core of papers about the golden road of OA, as well as its effects on citation and the impact of open access journals. Cluster 4 (blue color) refers to studies about the green road of OA, in which repositories and libraries occupy a central role. Cluster 5 (violet color) includes studies approaching the growth of OA in the directories of journals (DOAJ) and in repositories (for example, OpenDOAR), which are an expression of the level of development.

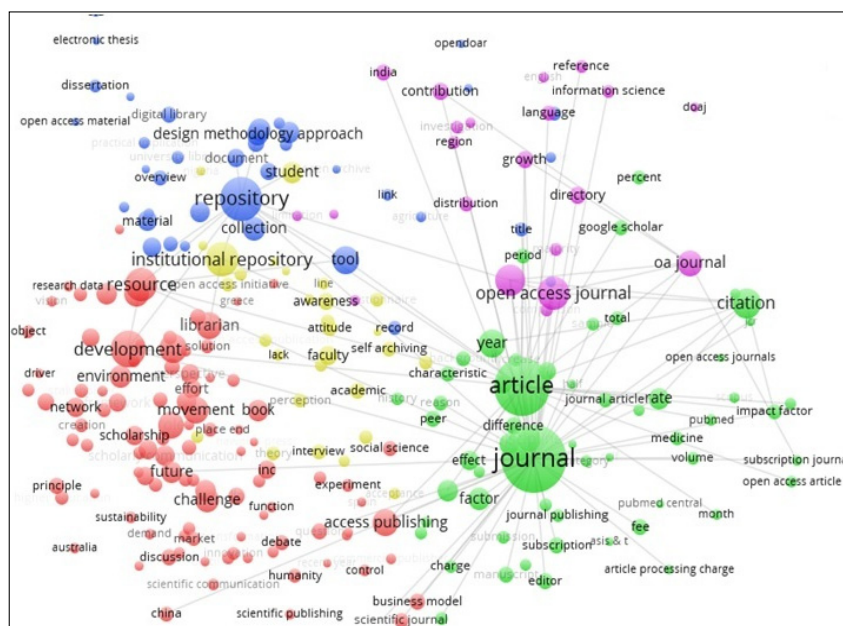


Figure 3. Net of the co-occurrence of words in titles and abstracts.

4. Final Remarks

This study aimed to characterize the most important moments of open access at the international level with respect to its presence in the world scientific literature, on a temporal analysis, to identify its scientific elite and core journals on the dissemination of new knowledge constructed on the theme. It also aimed to present the network of co-authorship highlighted on the theme and to identify the main subthemes and their co-occurrence in the analyzed literature, by keyword indexing of the articles.

It is important to notice that the open access movement considers knowledge as a good accessible to all citizens publicly in order to provide society with the new knowledge coming from new research.

Therefore, it is considered that for developing countries, this model is important to promote equity, as well as to increase the ability of researchers to advance in research through free access to science recently produced.

The important role of the conference convened by the Open Society Institute, in December 2011, in Budapest, with the subsequent declaration of BOAI, in which the principles, strategies, rules and commitments related to open access to research literature for an immediate increase in the scientific production on the theme were outlined, has been observed.

If, on the one hand, OA is relevant for distributing, in a democratic way, the new knowledge constructed, it has been found that the main countries producing on theme are: the United States, the United Kingdom, Spain and Germany, countries belonging to the science mainstream, but also countries, such as India, Brazil, Nigeria and Iran, among others, considered to be developing, occupy a prominent place in the ranking of the most productive countries in the theme.

This prominent place is paired with the most productive researchers, who come from countries, such as the United States, Canada, France and Spain, among others, such as the Russian Federation, India, Brazil and Malaysia, more outlying countries. Thus, according to these data, it is suggested that OA is important for countries in any development stage, as it proposes the growth and democratization of knowledge.

In relation to journals, the concern of authors to socialize the new knowledge generated on the theme in journals whose publication policy allows archive articles, either the green or golden road, has been noticed. It is highlighted that most journals belonging to the core have British origin (U.K.), belonging to Q1 and Q2. Besides, journals with significant impact on the academic community disseminate the new knowledge generated on the theme.

A collaborative network with a few subnets where co-authors are from different countries and several small subnets comprised of co-authors from the same country have been observed, showing that interlocation slightly surpasses the borders of their own countries. Perhaps, the recent theme growth and the policies not being fully disclosed explain the limited international co-authorship.

Clusters related to co-occurrences of keywords demonstrate five thematic groups of research: open access from the perspective of change in the model of scientific communication; the attitude and perception regarding OA by the academic community; the golden road of OA, as well as its effects on citation and the impact of journals with open access; the green road of OA, whose repositories and libraries occupy a central role; the growth of OA in journal directories (DOAJ) and the growth of repositories (for example OpenDoar), which are an expression of development acknowledgments.

This study, based on mainstream science literature, not only allows identifying the themes of debates that mark the development of open access at the international level, but also, from the analysis of the authors' profiles and the most prolific journals, it is possible to state that open access is one of the new emerging and frontier fields of library and information science.

We recognize, however, that several journals addressing OA as a subject are not yet indexed in Scopus. Thus, we suggest further studies be made using Google Scholar in order to visualize the behavior of the theme OA in journals and other type of academic documents that are not indexed in major international databases to have a broader and more complete visualization of the studied theme.

It is expected that this research has contributed to the advancement of knowledge about the theme and also that it encourages new studies to include aspects, such as the global identification of the areas in which open access has advanced more, the types of repositories adopted in the different countries and the institutions involved.

Author Contributions: The authors, Sandra Miguel, Ely Francina Tannuri de Oliveira and Maria Cláudia Cabrini Grácio, contributed equally to this paper, i.e., conceived, designed, analysed the data and wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Drott, M.C. Open access. *Annu. Rev. Inf. Sci. Technol.* **2006**, *40*, 79–109. [CrossRef]
2. White, S.; Creaser, C. *Trends in Scholarly Journal Prices 2000–2006*; LISU: Loughborough, UK, 2007. Available online: <http://www.lboro.ac.uk/microsites/infosci/lisu/downloads/op37.pdf> (accessed on 8 May 2010).
3. BOAI. *Budapest Open Access Initiative*; Open Society Institute: Budapest, Hungary. Available online: <http://www.soros.org/openaccess/read.shtml> (accessed on 21 september 2015).
4. Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. Available online: <http://openaccess.mpg.de/> (accessed on 15 September 2015).
5. Bethesda Statement on Open Access Publishing. Available online: <http://www.earlham.edu/~peters/fos/bethesda.htm> (accessed on 20 September 2015).
6. Björk, B.-C.; Laakso, M.; Welling, P.; Paetau, P. Anatomy of green open access. *J. Assoc. Inf. Sci. Technol.* **2014**, *65*, 237–250. [CrossRef]
7. Graziotin, D.; Wang, X.; Abrahamsson, P. A framework for systematic analysis of open access journals and its application in software engineering and information systems. *Scientometrics* **2014**, *101*, 1627–1656. [CrossRef]
8. Harnad, S.; Brody, T.; Vallieres, F.; Carr, L.; Hitchcock, S.; Gingras, Y.; Oppenheim, C.; Hajjem, C.; Hilf, E. The Access/Impact Problem and the Green and Gold Roads to Open Access: An Update. *Ser. Rev.* **2008**, *34*, 36–40. Available online: <http://eprints.ecs.soton.ac.uk/15852/> (accessed on 20 August 2015). [CrossRef]
9. Laakso, M. Green open access policies of scholarly publishers: A study of what, when, and where self-archiving is allowed. *Scientometrics* **2014**, *99*, 475–494. [CrossRef]
10. Rizor, S.L.; Holley, R.P. Open Access Goals Revisited: How Green and Gold Open Access Are Meeting (or Not) Their Original Goals. *J. Sch. Publ.* **2014**, *45*, 321. Available online: <http://dx.doi.org/10.3138/jsp.45.4.01> (accessed on 20 August 2015). [CrossRef]
11. Suber, P. Ensuring Open Access for Publicly Funded Research. *Br. Med. J.* **2012**, *345*. Available online: <http://dx.doi.org/10.1136/bmj.e5184> (accessed on 20 August 2012). [CrossRef] [PubMed]

12. Björk, B.-C.; Welling, P.; Laakso, M.; Majlender, P.; Hedlund, T.; Guðnason, G. Open access to the scientific journal literature: Situation 2009. *PLoS ONE* **2010**, *5*, e11273. [[CrossRef](#)] [[PubMed](#)]
13. Miguel, S.; Chinchilla-Rodríguez, Z.; Moya Anegón, F. Open access and Scopus: A new approach to scientific visibility from the standpoint of access. *J. Am. Soc. Inf. Sci. Technol.* **2011**, *62*, 1130–1145. [[CrossRef](#)]
14. Gargouri, Y.; Larivière, V.; Gingras, Y.; Carr, L.; Harnad, S. Green and Gold Open Access Percentages and Growth, by Discipline. Available online: <http://arxiv.org/abs/1206.3664> (accessed on 15 September 2012).
15. Düzyol, G.; Taşkın, Z.; Tonta, Y. Mapping the Intellectual Structure of the Open Access Field through Co-citation Analysis. In Proceedings of the IFLA Satellite Pre-Conference: Open Access to Science Information Trends, Models and Strategies for Libraries, Crete, Greece, 6–8 August 2010.
16. Craig, I.D.; Plume, A.M.; McVeigh, M.E.; Pringle, J.; Amin, M. Do open access articles have greater citation impact? A critical review of the literature. *J. Informetr.* **2007**, *1*, 239–248. [[CrossRef](#)]
17. Codina, L. Scopus: The greatest scientist of the web browser. *Prof. Inf.* **2005**, *14*, 44–49.
18. Fingerman, S. Scopus: Profusion and confusion. *Online* **2005**, *29*, 36–38.
19. Jacsó, P. As We May Search—Comparison of Major Features of the Web of Science, Scopus, and Google Scholar Citation-Based and Citation Enhanced Databases. *Curr. Sci.* **2005**, *89*, 1537–1547. Available online: <http://www.ias.ac.in/cursci/nov102005/1537.pdf> (accessed on 20 August 2015).
20. Fingerman, S. Web of Science and Scopus: Current Features and Capabilities. *Issues Sci. Technol. Librariansh.* **2006**, *48*. Available online: <http://www.istl.org/06-fall/electronic2.html> (accessed on 20 August 2015).
21. Jacsó, P. Comparison of journal impact rankings in the SCImago Journal & Country Rank and the Journal Citation Reports databases. *Online Inf. Rev.* **2010**, *34*, 642–657.
22. Bakkalbasi, N.; Bauer, K.; Glover, J.; Wang, L. Three options for citation tracking: Google Scholar, Scopus and Web of Science. *Biomed. Digit. Libr.* **2006**, *3*, 7. [[CrossRef](#)] [[PubMed](#)]
23. Falagas, M.E.; Pitsouni, E.I.; Malietzis, G.A.; Pappas, G. Comparison of PubMed, Scopus, Web of Science, and Google Scholar: Strengths and weaknesses. *FASEB J.* **2008**, *22*, 338–342. [[CrossRef](#)] [[PubMed](#)]
24. Braun, T.; Glänzel, W.; Schubert, A. How balanced is the Science Citation Index’s journal coverage? A preliminary overview of macro level statistical data. In *The Web of Knowledge—A Festschrift in Honor of Eugene Garfield*; Cronin, B., Atkins, H.B., Eds.; Information Today: Medford, NJ, USA, 2010; pp. 251–277.
25. Moya-Anegón, F.; Chinchilla-Rodríguez, Z.; Vargas-Quesada, B.; Corera-Álvarez, E.; Muñoz-Fernández, F.J.; González-Molina, A.; Herrero-Solana, V. Coverage analysis of Scopus: A journal metric approach. *Scientometrics* **2007**, *73*, 53–78. [[CrossRef](#)]
26. Karino, M.E.; Felli, V.E. A Enfermagem baseada em evidências: Avanços e inovações em revisões sistemáticas. *Cienc. Cuid. Saúde* **2012**, *11*, 11–15. [[CrossRef](#)]
27. Hagemann, M. Ten Years on, Researchers Embrace Open Access. *Voices*. Available online: <https://www.opensocietyfoundations.org/voices/ten-years-on-researchers-embrace-open-access> (accessed on 18 August 2015).
28. United Nations Educational, Scientific and Cultural Organization. Global Open Access Portal. The Netherlands. Available online: <http://www.unesco.org/new/en/communication-and-information/portals-and-platforms/goap/access-by-region/europe-and-north-america/netherlands/> (accessed on 20 September 2015).
29. Dekker, S. Open Access to Publication. Available online: <https://www.government.nl/documents/parliamentary-documents/2014/01/21/open-access-to-publications> (accessed on 11 May 2014).
30. Weitzel, S.R. As novas configurações do Acesso Aberto: Desafios e propostas. *Rev. Eletron. Comun. Inf. Inov. Saúde* **2014**, *6*, 65–75.
31. Glänzel, W. National characteristics in international scientific co-authorship. *Scientometrics* **2001**, *51*, 69–115. [[CrossRef](#)]
32. Person, O.; Glänzel, W.; Dannell, R. Inflationary bibliometric values: the role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics* **2004**, *60*, 421–432. [[CrossRef](#)]
33. Ennas, G.; di Guardo, M.C. Features of top-rated gold open access journals: An analysis of the scopus database. *J. Informetr.* **2015**, *9*, 79–89. [[CrossRef](#)]

