The MIT Press

Mapping the use of Google Scholar in evaluative bibliometric or scientometric studies: A bibliometric review

Fabiana Andrade Pereira [iD] and Rogério Mugnaini [iD]

University of São Paulo, São Paulo, Brazil

ABSTRACT

Google Scholar (GS) has aroused a good deal of interest among the bibliometric and scientometric community, owing to its capacity for gathering publication data, tracking citations, and creating metrics. This has led to reflections on its potential value as a means of enhancing evaluative procedures. However, despite being a useful tool because of its wide coverage, it has been monitored by specialists. For this reason, we aimed to map out the publications in the areas of Information Science & Library Science and/or Computer Science that make use of GS through a bibliometric review. Comprising data retrieved from the WoS and Dimensions, the results drew the attention of the bibliometric and scientometric community to the range of research problems in studies using GS. They also made it possible to identify the most prolific countries and authors and their preferred sources for publication. The presence of non-Anglophone countries and those from Latin America highlights the importance of alternative information sources to bibliometric and scientometric studies.

1. INTRODUCTION

Among the technical infrastructures available, information sources are an essential input for obtaining access to scientific publications and to assess the impact of scientific research. It is through these means that the process of evaluating scientific output can take place and allow indicators and analytical bibliometric tools to be made available. The sources of information generally used by the bibliometric and scientometric community have traditionally been databases and, more recently, academic search tools.

Gingras (2016) states that "the origins of the data that are used, represent a key factor in any kind of assessment." It has been confirmed that it is the responsibility of scientometric and bibliometric studies to discuss the importance of the data and information sources because it is through these means, combined with a high computational capacity for storage and rapid access to bibliographic data, that the evaluation of scientific output is made feasible. Thus, as Costas (2017) makes clear, there is a need to make use of indicators, tools, and applications for an analysis and understanding of both its internal infrastructure and external impacts. In this sense, it is important to differentiate between relational and evaluative bibliometric approaches (Thelwall, 2008). It is vital to consider both the source of data and its coverage. Any evaluative exercise will succeed better by understanding internal aspects of the disciplines, such as their cognitive structure and relationships, which can be analyzed via the literature.

Gusenbauer (2019) believes that the mechanisms of academic research, such as Google Scholar (GS) and Microsoft Academic, ensure a supply of robust information on scientific knowledge and various types of documents emerging from both formal and informal kinds of scientific communication. As well as offering greater accessibility, these also allow information filtering as an alternative way of retrieval.

According to Orduña-Malea, Martín-Martín, and Delgado López-Cózar (2017), since it was founded in 2004 GS has aroused a good deal of interest within the bibliometric and scientometric community. The authors believe that when being used as a tool, GS can be viewed from two standpoints: for discovery, or in other words as a search tool that can provide the user with a pleasant experience, owing to its wide-ranging capacity to explore scientific information quickly and easily; and its usefulness in assessing academic achievements. The second factor is given prominence because of the increasing reliance on GS by users and professionals as a bibliometric tool for various evaluative purposes, as it makes it possible to assess the impact of documents and the authors who publish them.

Delgado López-Cózar, Orduña-Malea, and Martín-Martín (2019) add that owing to its inherent features, GS has triggered the beginning of a revolution in the marketplace of scientific information. As well as being easy to use, it also has a wide coverage, and is able to automatically and rapidly incorporate academic information available on the web in the index, unlike the traditional databases Web of Science (WoS) and Scopus. Moreover, the search tool shows the number of citations, as well as the subsequent development of secondary products such as GS Citation (GSC) and GS Metrics (GSM). Nonetheless, despite all these attributes, many people are questioning its potential value as a data source for bibliometric analyses.

In the case of Latin America and the Caribbean (LA&C), in particular, GS plays a different role because, owing to its wide coverage, it is able to access and analyze the impact of journals on these countries, as they are often available in institutional repositories and digital libraries, such as SciELO. As is pointed out by Canto, Pinto et al. (2022), it is a way of drawing attention to the journals of a region that have never featured prominently in the scientific field among the international databases. Although these are often concerned with topical issues affecting particular areas and written in regional languages, they now have an open access publication model, and their impact must be measured considering regional sources (Santos, Fraumann et al., 2021).

In view of the worldwide popularity of GS as a source of scientific information—which perhaps explains the adoption of its metrics for evaluative purposes—there is an increasing need to understand how and for whom it is being used. In addressing the questions related to measuring scientific information in the context of a general assessment, the aim of this article is to map out the studies that make use of this research tool by embarking on a bibliometric analysis concerned with this issue. It is hoped that this will contribute to the debate about its wide-ranging use for assessing and measuring scientific achievements in various countries, while at the same time adding to previous literature review studies on GS.

In the current debate about the data sources for scientometric studies, the bibliometric and scientometric community is increasingly resorting to GS. Its potential value as a tool that can assist in giving exposure to scientific output is well known, but its limitations suggest that the community is aware of the possible implications of using it extensively when making an assessment of scientific performance. Thus, this article seeks to estimate the extent to which scientific output mentions the GS in bibliometric and scientometric studies and give thought to the effects of GS on measuring and assessing scientific information.

2. METHODOLOGY

The data for this study consist of publications on various document typologies that mention GS, and which can be retrieved from the WoS and Dimensions databases. The choice of WoS can be explained by the fact that it has many specialized sources in its Core Collection that are often used by the bibliometric and scientometric community, and the choice of Dimensions is due to the fact that it complements WoS, with greater coverage, which is expected to be more representative for Latin America. In addition, because of the need to clean data from both sources, the features offered by Dimensions to search and export the data were more prevalent than the use of GS, in addition to the evidence that its coverage is greater for recent literature (Orduña-Malea & Delgado-López-Cózar, 2018).

A preliminary bibliometric analysis made it clear that some of the features of the literature in such a subject, as could be expected, show a greater prevalence of documents published in Information Science & Library Science (IS&LS) areas and/or Computer Science. This is verified mainly when considering the output of the most prolific authors, renowned specialists among the bibliometric and scientometric community. Hence, the corpus of this study was restricted to documents published in sources from these areas of classification.

Furthermore, literature review studies appeared more frequently in WoS in Health Science areas (around 80% of the documents), and much less frequently in IS&LS and Computer Science (16.3%). Health Science areas are increasingly using both bibliometric methods and GS as a data source, bringing a significant number of articles that fall outside the scope of this study. In parallel, the second most frequent type of document in Dimensions is the preprint, representing 4.5% in Health Science areas, but 10.2% in IS&LS and Computer Science. These facts highlight the complementarity of each data source, owing to the different kinds of document types found in each one.

The data were gathered on June 20, 2022 from both data sources. The following terms were used in the search strategy ["google scholar" and ("bibliometr*" or "cientometr*" or "scientometr*" or "evaluat*" or "assessm*")], which had to occur in Title, Abstract, or Author Keywords.

In WoS, the records were restricted to "Information Science Library Science or Computer Science" in the field Research Areas, without any period limit, resulting in 433 documents. Dimensions returned 573 documents from sources classified in the "08 Information and Computing Sciences" Field of Research.

The documents of each data source were then analyzed to exclude preprints and proceedings papers that were subsequently published in journals, maintaining the same title. This action reduced the documents from Dimensions to 534 documents.

The data from the two sources were compared, resulting in an overlap of 203 documents, with 230 exclusively from WoS and 331 exclusively from Dimensions, giving 764 documents.

As mentioned above, studies that mention GS as an information source for literature review, but which were not related to the scope of this study, were excluded. The coverage of GS attracted attention in many areas, resulting in Gehanno, Rollin, and Darmoni (2013) testing its suitability to be used as unique source in systematic review. Sometimes the number of citations is used to set the corpus of the studies or even to perform "bibliometric and contents analyses based on a literature review" (Gómez-Gil, Flo et al., 2020).

Documents that mention the following terms in title or abstract were considered to be excluded: "scoping review" OR "bibliometric review" OR "bibliographic anal" OR "systematic review" OR "literature review", as were those that mention these terms in the abstract:

"comprehensive review" OR "narrative review" OR "evidence-based review" OR "review study" OR "topical Assessment" OR "state-of-". After reading each abstract, 245 documents were discarded and the final corpus comprised 519 items, with the following distribution between the data sources: 168 from WoS, 155 from both sources, and 196 from Dimensions.

The bibliographic fields selected for the definition of the variables in the study were as follows: data source, year of publication, type of document, open or closed access to the document, publication source, author, and author affiliation country.

3. RESULTS AND DISCUSSION

The temporal distribution of the scientific output that refers to GS shows that in 2017 the number of articles published began to be significantly higher (Figure 1). The average number of documents during this period rose from 15.8 to 54.8, which represents a growth of 247%.

However, when account is taken of the different data sources, it is clear that about 32.4% can only be found in WoS, 37.8% are only in Dimensions and the remainder (29.9%) exist in both sources. As can be seen, Dimensions supplements WoS in a significant way, as it enables a broader scene to be laid out of references to GS in the whole output of the world.

When the different types of documents are analyzed (Table 1), the original articles recur more often—independently of the source. Proceedings papers are the second most common type in WoS, although they make way for "preprints" when they are regarded as the only documents that can be found in Dimensions. The review articles are found mainly in WoS, with half of them also in Dimensions. With regard to other documents, those exclusively from Dimensions are book chapters, and in WoS letters are predominant, followed by editorial material.

The percentage of documents in open access can be more clearly seen in the documents that are found only in Dimensions (67.3%), whereas in WoS (with and without Dimensions) it is about 48.3%—preprints are an especially important factor in explaining this difference. In the same way, when the original articles are included, a significantly higher percentage can be seen for Dimensions (70.0%), which underlines their capacity for picking out the literature

Figure 1. Number of documents per data source and year of publication and percentage of documents exclusively in Dimensions.

Table 1. Number of documents, according to type, data source, and percentage of documents in open access

Document type		
Source	#documents	Open access (%)
Article	**398**	**61.1**
Dimensions	150	70.0
WoS	128	53.1
WoS/Dimensions	120	58.3
Proceedings paper	**66**	**13.6**
WoS	32	3.1
WoS/Dimensions	21	23.8
Dimensions	13	23.1
Article; Proceedings paper	**5**	**40.0**
WoS/Dimensions	3	33.3
WoS	2	50.0
Review	**10**	**60.0**
WoS/Dimensions	5	60.0
WoS	5	60.0
Preprint	**21**	**100.0**
Dimensions	21	100.0
Other	**19**	**36.8**
Dimensions	12	25.0
WoS/Dimensions	6	50.0
WoS	1	100.0
Total	**519**	**55.5**

available in open access. The same can be found with regard to proceedings papers, which, despite having a lower percentage in open access, show a significant difference when account is taken of the documents that are only found in WoS.

Table 2 shows the publication sources with at least 1.5% of the documents in each data source. The total number of sources of the corpus is 285 titles, with 162 exclusively in Dimensions and 128 in WoS (with and without Dimensions). In WoS, it can be seen that the 12 positions are filled by 11 journals and a conference proceedings that published 52.9% of the documents, of which *Scientometrics* accounts for the highest percentage. This is followed by *Journal of Informetrics*, which is in second place, and *JASIST* (*Journal of the American Society for Information Science and Technology*). Two other journals follow: *Profesional de la Información* and *Malaysian Journal of Library & Information Science*.

Table 2. The most widely used publication sources in each data source

Publication source (WoS and WoS/Dimensions)	WoS (total)	
	#	%
Scientometrics	68	21.1
Journal of Informetrics	22	6.8
JASIST	13	3.1
Profesional de la Informacion	12	4.0
ISSI Conference	10	3.7
Malaysian Journal of Library & Information Science	8	2.5
Online Information Review	8	2.5
Global Knowledge Memory and Communication	7	2.2
Revista Espanola de Documentacion Cientifica	7	2.2
Publications	6	1.9
IEEE Access	5	1.5
Journal of Scientometric Research	5	1.5
Other 116	152	47.1

Publication source (Dimensions)	Dimensions	
	#	%
arXiv	8	4.1
SSRN Electronic Journal	4	2.0
Lecture Notes in Computer Science	4	2.0
Evidence Based Library and Information Practice	3	1.5
Innovation in Teaching and Learning in Information and Computer Sciences	3	1.5
PLOS ONE	3	1.5
Information Technologies and Learning Tools	3	1.5
JMIR Preprints	3	1.5
Other 155	165	84.2

In the whole data, proceedings papers come from a variety of different events, but just two present more than one document—even considering all editions together. The most frequent is the *International Society for Informetrics and Scientometrics* (ISSI), whose importance for the bibliometric and scientometric community is well known (Fraumann, Mugnaini, & Sanz-Casado, 2021). The proceedings of the editions in 2015, 2017, and 2019 presented 18, 17, and 34 articles respectively—taking account of the occurrence of the term "Google Scholar" in the full text. This highlights the attention of this community to it. However, in the present study 13 documents were
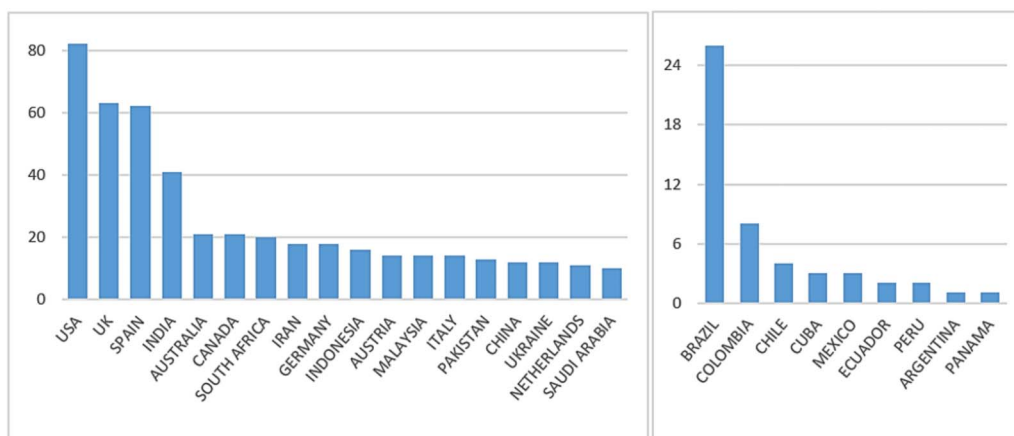
Figure 2. Number of documents per affiliation country of authors from (left panel) most prolific countries outside LA&C and (right panel) LA&C countries.

identified in WoS (Table 2) as being one of those also identified in Dimensions. Finally, the *International Conference on World Wide Web* (WWW) presented two documents.

It is worth noting the following repositories: *arXiv, SSRN, Electronic Journal, JMIR Preprints* and the chapters of the book series *Lecture Notes in Computer Science*. In addition, there are some local journals, as well as *PLOS ONE*.

Figure 2 shows that the countries with the most publications are the United States, United Kingdom, Spain, and India (left panel), because their output derives from a significantly larger number of authors: 171, 112, 70, and 99 respectively. Brazil (with 88 authors) is in fifth place in the general ranking but the leading country in Latin America & Caribbean (right panel), followed by Colombia and Chile.

One interesting factor related to the ranking concerns the position of non-Anglophone countries, such as Spain and Brazil, as they do not usually have their scientific output well represented in databases such as WoS. For this reason, Australia and Canada (more prolific countries in the lingua franca) lose their position because the linguistic bias justifies non-Anglophone countries using alternative citation indices to capture regional citations in bibliometric and scientometric studies (Santos et al., 2021). To reinforce this tendency, it is important to highlight the specialist community of Brazil, which in 2007 showed more significant growth of its production in bibliometric and scientometric studies in GS than in WoS (Meneghini & Packer, 2010).

GS serves as an important data source for researchers, as well as professionals in the area of bibliometrics and scientometrics, because there is a need to give a minimum amount of information about the output of authors who have at least three articles in the corpus of this study (Table 3).

The most prolific author in the United States is Peter Jacsó (University of Hawaii), with 15 articles, who is known for his articles that provide an exhaustive list of errors that can be found in GS. Sometimes his work includes comparisons with other data sources, is highly critical of the tools, and issues a warning about those who defend them as entirely suitable for making an appraisal of scientific information.

Following him is William H. Walters, the current executive director of Manhattan College, who has five articles. His work on GS covers several disciplinary fields and addresses the

Table 3. Authors from the most prolific countries outside LA&C with at least three documents

Country	#	Main data source	Author	Institution
United States	11	WoS	Jacso, Peter	University of Hawaii
	5	WoS	Walters, William H.	Millersville University, Pennsylvania
	5	WoS	Meho, Lokman I.	Indiana University
	4	WoS	Yang, Kiduk	Indiana University
United Kingdom	16	WoS	Thelwall, Mike	University of Wolverhampton
	8	WoS	Kousha, Kayvan	University of Wolverhampton
	6	WoS	Harzing, Anne Wil	Middlesex University
	4	WoS	Mingers, John	University of Kent
	3	WoS	Willett, Peter	University of Sheffield
Spain	22	WoS	Delgado Lopez Cozar, Emilio	University of Granada
	18	WoS	Orduña-Malea, Enrique	Polytechnic University of Valencia
	14	WoS	Martín-Martín, Alberto	University of Granada
	8	WoS	Ortega, Jose Luis	CSIC, Spain
	6	WoS	Aguillo, Isidro F.	CSIC, Spain
	5	WoS	Torres Salinas, Daniel	University of Navarra
	4	WoS = Dimensions	Ayllon, Juan Manuel	University of Granada
	4	WoS	Cabezas Clavijo, Alvaro	University of Granada
	4	WoS	Jimenez Contreras, Evaristo	University of Granada
India	3	WoS	Garg, Kailash C.	NISTADS
Australia	4	WoS	Alakangas, Satu	University of Melbourne
	4	WoS	Harzing, Anne Wil	Middlesex University
Canada	4	WoS	Serenko, Alexander	Lakehead University
Germany	4	WoS	Bornmann, Lutz	Max Planck Society
	3	WoS	Dorsch, Isabelle	Heinrich Heine University Dusseldorf
South Africa	6	WoS	Onyancha, Omwoyo Bosire	University of South Africa
Malaysia	5	WoS	Zainab, A. N.	University of Malaya
Austria	6	WoS	Gorraiz, Juan	University of Vienna
	4	WoS	Gumpenberger, Christian	University of Vienna
	3	WoS	Wieland, Martin	University of Vienna
Norway	3	WoS	Gjesdal, Øyvind Liland	University of Bergen
	3	WoS	Mikki, Susanne	University of Bergen
Israel	3	WoS	Bar-Ilan, Judit	Bar Ilan University

question of performance and the accuracy of research in GS, as well as making a comparative assessment of other academic data sources.

Lokman Meho, who is affiliated with Indiana University, also has five articles (and another one signed as a Lebanese affiliation). His studies of GS are concerned with conducting citation analysis, and particularly stress the need to merge bibliometric methods with data sources to ensure a multidimensional evaluative approach that is less biased. His studies also show his concern with measuring the research performance of research staff and institutions, as well as assessing the impact they have on decision-making and planning of research policies and practices. Kiduk Yang (Indiana University) contributed to four of his articles.

From the United Kingdom, Mike Thelwall (University of Wolverhampton) appears as the most prolific author, with 26 articles. He forms a part of the Statistical Cybermetrics and Research Evaluation Group and is known for being devoted to altmetric studies, webometrics, metrics in the social web, and the analysis of feelings. Kayvan Kousha is a frequent collaborator, who despite working in the same institution, also signs his name on articles with an Iranian affiliation. These two have developed collection methods and an analytical system that has an academic impact on research, and is outside the traditional citation indices, yet involving different types of web data such as GS, Google Books, Google Patents, Microsoft Academic, and Wikipedia. Together with the three most prolific authors from Spain, Kousha has signed his name on two studies with macrolevel approaches, both aimed at comparing the coverage of GS citations with other sources. This is evidence of the potential value of forming a collaboration between different groups—that of 2018 compared with the WoS and Scopus and that of 2021 with the addition of Microsoft Academic, Dimensions, WoS, and OpenCitations' COCI.

Professor Anne-Wil Harzing (Middlesex University) has six articles, and also features in Table 3 as a researcher Australia. Since 2006, she has made a contribution to bibliometric studies throughout the world by having proposed Publish or Perish software. This free software allows data collection for the publication and citation of GS by means of search expression, and includes impact metrics together with the *h*-index. Thus, to a great extent, the works of the author reflect the use of GS (from whence the information is extracted), by devoting research to making comparative analyses with traditional databases in a macrolevel approach. At the same time, the author works with the three most prolific authors from Spain, which is evidence of their centralization in collaborations on the subject.

John Mingers (University of Kent) submitted four articles to the corpus, three of which are worth mentioning, as their methodology is aimed at normalizing GS citations for the purposes of evaluative bibliometric analyses at various levels, although it should be noted that the data from this source are always less reliable.

Peter Willett (emeritus professor at the University of Sheffield) was the author of five articles between 2008 and 2011. Three of them are signed by Aryati Bakri (from the same university, who also signs his name to work with an affiliation to Malaysia, his country of origin). Together they employed GS for data collection for the publication and citation of two journals in Malaysia and for several researchers in the computing departments of universities in the same country. Working by himself, Peter Willett was the author of a study of the websites of UK departments of library and information science, in which he analyzed the correlation between webometric and bibliometric indicators (those based on citations obtained from GS). In his other article, he also focused on the area of IS&LS and worked together with Michael Norris and Charles Oppenheim (both from Loughborough University), comparing the citation indicators and expert judgments on research published by 101 scholars.

Table 3 shows that Spain has the largest number of authors, belonging to four institutions, prominent among them being the University of Granada, to which five of them belong. Regular collaborators Emílio Delgado-López-Cózar, Enrique Orduña-Malea, and Alberto Martín-Martín perform research using GS and are widely recognized in the scientific community. Their studies include GS coverage (considering documents and citations, research fields, documentary typology, authors, languages, and comparisons with other data sources), updating of data, data errors and limitations, and carrying out searches, as well as the use of GS for assessment purposes (of researchers and journals). Álvaro Cabezas-Clavijo and Juan Manuel Ayllón are members of the group, signing their names to some of the documents.

José Luis Ortega and Isidro F. Aguillo, both members of the *Consejo Superior de Investigaciones Científicas*, collaborated on four articles, the oldest being written in 2010 and setting out a proposal for the Ranking Web of World Repositories. In establishing this, the authors resorted to GS to quantify the volume of PDF files and number of items in the repositories. In the years that followed, they published three other articles, based on profiles of authors taken from GSC: to find information about their affiliations; to map out the keywords; and to conduct a comparative analysis of the documents and citations with Microsoft Academic Search. José Luis Ortega provided five other articles of which he was the author himself and where it was clear he had adopted an evaluative approach. This involved the following: the relationship between bibliometrics and altmetrics indicators; peer-review activities and bibliometric performance; and different analytical methods, such as coauthorship networks, decision trees, and a longitudinal demographic study of the population of GSC. Isidro wrote an article by himself which adopted a webometric approach to conduct an analysis of institutional web domains; and another, in coauthorship with the research team of the English professor Mike Thelwall, who analyzed the use of the web and social websites by highly cited researchers.

Evaristo Jiménez-Contreras and Daniel Torres-Salinas have articles in common that are worth highlighting: a comparison between the number of times monographs were lent from libraries and the number of citations and the assessment of a bibliometric mobile application, which makes it possible to analyze the rankings of researchers at the University of Granada, with the aid of data from GS Profiles.

From India, Kailash C. Garg (of the National Institute of Science, Technology and Development Studies) has three articles, which make use of GS as an alternative means of carrying out bibliometric studies about scientific areas, articles by Indian authors, and issues of regional interest.

In the case of Australia, Satu Alakangas (University of Melbourne) should be mentioned because he has worked together with Anne-Wil Harzing and published four articles. The work is concerned with the coverage of GS and how it compares with other databases and the *h*-index; in addition, these authors are responsible for an important revision of Microsoft Academic.

From Canada, there is Alexander Serenko (currently a professor at Ontario Technical University); his work on GS is mainly concerned with the analysis of citations from periodicals regarding knowledge management or artificial intelligence.

Finally, from Germany, Lutz Bornmann (Max Planck Society) presents four articles. He carried out a wide range of empirical studies with data from GS (including normalization of citations and correlation between citations and information obtained from peer review, as well as putting forward a scheme for a "meta-ranking" of journals in the area of Economics). Isabelle Dorsch (University of Dusseldorf), who also has three articles, analyzes the exposure of authors to various sources, including GS.

Omwoyo Bosire Onyancha is a prolific author from South Africa who uses GS data to analyze African journals and universities, as well as the impact of theses; studies on the assessment of South African researchers; and one article analyzing the terminology of indigenous knowledge.

Awang Ngah Zainab comes from Malaysia and analyzes the impact of journals as well as the citations obtained from items in an open access database (*Malaysian Abstracting and Indexing*—MyAis), and regularly publishes articles in a journal in his country (as noted in Table 2).

Juan Gorraiz, Christian Gumpenberger, and Martin Wieland come from Austria and are coauthors of articles. They make use of data from GS for collection of publications and citations for an analysis of sources in the area of Geography, and about the work of some artistic and scientific celebrities. They also collaborate with some of the Spanish authors highlighted in Table 3.

Øyvind Liland Gjesdal and Susanne Mikki, from Norway, collaborate in two articles, performing different studies focused on the open availability of articles. They use GS to collect publications and citations, using the national Norwegian scientific output (Cristin, the Current Information System in Norway) as the reference data set.

Judit Bar-Ilan was a notable figure from Israel whose studies assessed both the drawbacks and benefits of GS, as well as writing an extensive review in an article about GS and its ability to supply data for scientific evaluation. It is also worth mentioning a recent article, published in 2020 by Gali Halevi (from the United States), that consists of a posthumous homage to the legacy of Judit Bar-Ilan.

With regard to Latin America and the Caribbean (Table 4), Brazil is the country with most authors. Fábio Lorensi do Canto (Federal University of Santa Catarina) has published work in partnership with Adilson Luiz Pinto, Edson Mário Gavron (from the same university), and Marcos Talau (Federal Technological University of Paraná). Their work involving GS is concerned with conducting an analysis of Brazilian periodicals and/or Latin-American and Caribbean articles indexed in GSM. However, one of them is restricted to Brazilian journals from IS&LS, and employs the methodology put forward by the new Qualis (a classification system of journals that is used for assessment purposes in the national arena).

Among the Colombian researchers, Alejandro Uribe-Tirado (Universidad de Antioquia) has two studies. One combines bibliometric and altmetric data from various sources, including

Table 4. Authors from the LA&C countries with at least two documents

Country	#	Main data source	Author	Institution
Brazil	3	WoS	Canto, Fábio Lorensi do	Federal University of Santa Catarina
	3	WoS	Pinto, Adilson Luiz	Federal University of Santa Catarina
	2	WoS	Gavron, Edson Mário	Federal University of Santa Catarina
	2	WoS	Talau, Marcos	Federal Technological University of Paraná
Colombia	2	WoS = Dimensions	Uribe-Tirado, Alejandro	Universidad de Antioquia
Cuba	2	WoS	Céspedes-Villegas, Alejandro	Marta Abreu Central University of Las Villas
	2	WoS	Paz-Enrique, Luis Ernesto	Marta Abreu Central University of Las Villas

GS, to compare Ibero-America researchers with publications on the subject of Information Literacy. In another article with Colombian colleagues, a detailed analysis of a national journal of librarianship was carried out, using bibliometric and altimetric indicators from various data sources, including GS.

Finally, in Cuba there are Alejandro Céspedes-Villegas and Luis E. Paz-Enrique (both from the Marta Abreu Central University of Las Villas). Their work, which uses GS, stresses the current need for alternative indicators to measure the scientific output of this Cuban university in social and academic networks. Because the question of the exposure of the scientific impact must be taken into account, a key factor is the presence and establishment of scientific communities in social networks and Web 2.0 platforms. For this reason, they seek to calculate the indicators and conduct comparative analyses within academic social networking sites, such as ResearchGate, among others.

Among the authors listed in Tables 3 and 4, the Colombian Alejandro Uribe-Tirado and the Spanish Juan Manuel Ayllón are the only ones whose output had the same number of articles in each data source. Thus, it can be seen that despite representing an important volume of documents for analysis, Dimensions did not prove to be important for the individual analysis of the most prolific authors. This suggests that their sources of sole publication have not been repeatedly used by authors that published under GS subjects in the areas of IS&LS and Computer Science.

4. FINAL REMARKS

This study illustrates how there has been a significant rise in the number of bibliometric and scientometric studies that have referred to GS in recent years, as mapped out in the WoS and Dimensions data sources. The combined use of databases reveals their supplementary benefits when it is taken into account that although WoS concentrates on specialist sources in the areas of IS&LS and/or Computer Science, Dimensions abounds in various sources, in particular preprints and proceedings. The prevalence of open access documents was observed in Dimensions, and this is an advantage of the collection of data in this source as well.

Another factor that should be stressed is the position of non-Anglophone countries among the more prolific in output. The position of Spain and Brazil in the global ranking, as well as Asian countries, is worth mentioning, as is the presence of the other two Latin American countries whose authors took part with the specialists. This signals the importance that GS has for analysis carried out by authors whose scientific production has offered several diagnoses on the use of this data source.

The output of some prolific authors in the subject, and their collaborators, has led to a valuable collection of articles, which reveal the following: the sheer size and documentary diversity of GS; its potential value for providing material that can be used for citation analysis; and various limitations in this source, as it makes a number of significant errors in its metadata when tracking the web in search of scientific information. It is worth noting the Publish or Perish software, because researchers in general recognize its usefulness.

This bibliometric review highlights the types of data and/or prevailing information showing that the publication and citation data (considered together) are frequently used in the methodology of studies, followed by citations, publications, and metrics, although the monitoring of the literature by specialists covers a broader wide array of research problems.

It can be concluded from the particular features noted in these studies, and in particular the limitations described by several of them, that guidance should be offered to the following:

those that plan their own methodological proceedings; those that make use of available metrics services; and, mainly, the assessment exercises that make use of GS and the resulting indicators for the determination of objective criteria.

REFERENCES

Canto, F. L., Pinto, A. L., Gavron, E. M., & Talau, M. (2022). Latin American and Caribbean journals indexed in Google Scholar Metrics. *Scientometrics*, *127*, 763–783. https://doi.org/10.1007/s11192-021-04237-x

Costas, R. (2017). Discussões gerais sobre as características mais relevantes de infraestruturas de pesquisa para a Cientometria. In R. Mugnaini, A. Fujino, & N. Y. Kobashi (Eds.), *Bibliometria e cientometria no Brasil: Infraestrutura para avaliação da pesquisa científica na Era do Big Data* (pp. 43–66). São Paulo: ECA/USP.

Delgado López-Cózar, E., Orduña-Malea, E., & Martín-Martín, A. (2019). Google Scholar as a data source for research assessment. In W. Glänzel, H. F. Moed, U. Schmoch, & M. Thelwall (Eds.), *Springer handbook of science and technology indicators* (pp. 96–126). New York: Springer. https://doi.org/10.1007/978-3-030-02511-3_4

Fraumann, G., Mugnaini, R., & Sanz-Casado, E. (2021). International conferences of bibliometrics. In R. Ball (Ed.). *Handbook bibliometrics* (pp. 65–73). Berlin, Boston: De Gruyter Saur. https://doi.org/10.1515/9783110646610-008

Gehanno, J. F., Rollin, L., & Darmoni, S. (2013). Is the coverage of Google Scholar enough to be used alone for systematic reviews? *BMC Medical Informatics and Decision Making*, *13*(1), 7. https://doi.org/10.1186/1472-6947-13-7, PubMed: 23302542

Gingras, Y. (2016). *Os desvios da avaliação da pesquisa: O bom uso da bibliometria.* Rio de Janeiro: UFRJ.

Gómez-Gil, E., Flo, M., Fernández, R., Esteva, I., & Gómez-Gil, F. J. (2020). Spanish research in gender dysphoria: A review of more than 20 years of biomedical literature. *Actas Españolas de Psiquiatría*, *48*(6), 271–286. https://www.actaspsiquiatria.es/repositorio//22/128/ENG/22-128-ENG-271-286-169398.pdf

Gusenbauer, M. (2019). Google Scholar to overshadow them all? Comparing the sizes of 12 academic search engines and bibliographic databases. *Scientometrics*, *118*, 177–214. https://doi.org/10.1007/s11192-018-2958-5

Meneghini, R., & Packer, A. L. (2010). The extent of multidisciplinary authorship of articles on scientometrics and bibliometrics in Brazil. *Interciencia*, *35*(7), 510–514. https://www.redalyc.org/pdf/339/33914381007.pdf

Orduña-Malea, E., & Delgado-López-Cózar, E. (2018). Dimensions: Re-discovering the ecosystem of scientific information. *El Profesional de la Información*, *27*(2), 420–431. https://doi.org/10.3145/epi.2018.mar.21

Orduña-Malea, E., Martín-Martín, A., & Delgado López-Cózar, E. (2017). Google Scholar como una fuente de evaluación científica: Una revisión bibliográfica sobre errores de la base de datos. *Revista Española de Documentación Científica*, *40*(4), 1–33. https://doi.org/10.3989/redc.2017.4.1500

Pereira, F. A., & Mugnaini, R. (2022). Data for "Mapping the use of Google Scholar in evaluative bibliometric or scientometric studies: A bibliometric review." *Zenodo*. https://doi.org/10.5281/zenodo.7439216

Santos, S. M. dos, Fraumann, G., Belli, S., & Mugnaini, R. (2021). The relationship between the language of scientific publication and its impact in the field of public and collective health. *Journal of Scientometric Research*, *10*(1s), s78–s87. https://doi.org/10.5530/jscires.10.1s.24

Thelwall, M. (2008). Bibliometrics to webometrics. *Journal of Information Science*, *34*(4), 605–621. https://doi.org/10.1177/0165551507087238