BEST PRACTICES IN BIBLIOMETRICS & BIBLIOMETRIC SERVICES

EDITED BY: Juan Ignacio Gorraiz, Rafael Repiso, Nicola De Bellis and Gernot Deinzer

PUBLISHED IN: Frontiers in Research Metrics and Analytics

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88971-969-3
DOI 10.3389/978-2-88971-969-3

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.
By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

BEST PRACTICES IN BIBLIOMETRICS & BIBLIOMETRIC SERVICES

Topic Editors:
Juan Ignacio Gorraiz, University of Vienna, Austria
Rafael Repiso, Universidad Internacional De La Rioja, Spain
Nicola De Bellis, University of Modena and Reggio Emilia, Italy
Gernot Deinzer, University of Regensburg, Germany

Citation: Gorraiz, J. I., Repiso, R., De Bellis, N., Deinzer, G., eds. (2021). Best Practices in Bibliometrics & Bibliometric Services. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88971-969-3

Table of Contents

frontiers in Research Metrics and Analytics

EDITORIAL
published: 11 November 2021
doi: 10.3389/frma.2021.771999

Editorial: Best Practices in Bibliometrics & Bibliometric Services

*Juan Ignacio Gorraiz**

Department of Bibliometrics and Publication Strategies, University of Vienna, Vienna, Austria

Keywords: bibliometrics, scientometrics, bibliometric services, best praclices, informetrics

Editorial on the Research Topic

Best Practices in Bibliometrics & Bibliometric Services

From my point of view, sciences are very similar to languages. Just as one can speak of dead and living languages, this also applies to the sciences in general and to bibliometrics and scientometrics in particular.

Pritchard already defined bibliometrics as "the application of mathematics and statistical methods to books and other media of communication" in order to "shed light on the processes of written communication and of the nature and course of development of a discipline".

However, most scientometric journals focus on publishing articles dealing with the introduction of new indicators, the exploration of new methodological techniques, the analysis of new instruments and data sources or the collection and comparison of the results traced from different tools. Contributions of a practical nature showing best practices in different institutions, discussing responsible and sound use of the different metrics, or suggesting new and innovative services for scientists, the administration and science policy makers, are usually rejected despite being of high interest. The reason for the rejection is that they do not contain novel or original research results.

This generates a tendency to favour those scientists who work in their ivory towers and publish an endless number of works without practical use, to the detriment of those ones working from a more practical way, trying to apply correctly indicators and methods, revealing and learning from their deficiencies, and refining and adapting them to suit the needs of the different target groups.

Predominance of theoretical publications makes scientometrics a "dead" discipline, in very clear contradiction with its genuine definition according to Pritchard. A research field is like a language, if it does not find application, it dies. Current research on bibliometrics does not respond to professional needs appropriately. Of course, it should also not only respond to professional needs. Without a solid and innovative theoretical background, we could never build a new discipline and achieve any goal. But, I think that we should also not run the risk of converting bibliometrics in a dead discipline.

To this purpose, it is necessary to bridge the gap between research and professionals conducting bibliometric analyses. We should not forget that science policy and librarian are usually the ones in charge of bibliometric analysis and that, for this reason, their contribution to the discourse is of great importance, as they are best placed to detect problems, benefits and shortcomings in the application of theoretical concepts. But, why is the community still reluctant considering librarians as researchers? Is not "Library and Information sciences" just another discipline more, like religion, politics, economics, or computer sciences?

On the other side, the lack of published examples of practical applications contrasts with the growing number of manifests and recommendations (e.g. San Francisco Declaration on Research Assessment (DORA), Leiden Manifesto, or more recently, the Honk Kong

Edited and reviwed by:
Zaida Chinchilla-Rodríguez,
Consejo Superior de Investigaciones Científicas (CSIC), Spain

****Correspondence:***
Juan Ignacio Gorraiz
juan.gorraiz@univie.ac.at

Specialty section:
This article was submitted to Research Assessment, a section of the journal Frontiers in Research Metrics and Analytics

Received: *07 September 2021*
Accepted: *26 October 2021*
Published: *11 November 2021*

Citation:
Gorraiz JI (2021) Editorial: Best Practices in Bibliometrics & Bibliometric Services. Front. Res. Metr. Anal. 6:771999. doi: 10.3389/frma.2021.771999

Principles, etc.,) that appear constantly and underlines the need to seek best practices and curb misuse.

However, these initiatives are generally reduced to prevent misuse or give recommendations, instead of providing practical guidance. Therefore, we need concrete examples of responsible use of bibliometrics to be published in order to revive, reinforce and refresh this young discipline.

The purpose of this Research Topic was to gather critical contributions from researchers who are able to share their experiences, initiatives, projects, policies or other insights concerning best practices in bibliometrics. Thus, it provides a short compilation of original applied bibliometric knowledge at the micro-, meso- and macro-level, as well as the description of responsible and innovative bibliometric services. It will also help to refrain from bad practices that are affecting the development of this discipline and contributing to its discredit.

Finally, I would like to thank all the authors for their collaboration and dedication, which was not easy to obtain.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Research Metrics and Analytics

ORIGINAL RESEARCH
published: 08 September 2020
doi: 10.3389/frma.2020.00005

Tracing Long-Term Outcomes of Basic Research Using Citation Networks

James Onken[1]*, Andrew C. Miklos[2] and Richard Aragon[2]

[1] Research Enterprise Analytics, LLC, Rockville, MD, United States, [2] Division of Data Integration, Modeling, and Analytics, National Institute of General Medical Sciences, Bethesda, MD, United States

In recent years, the science of science policy has been facilitated by the greater availability of and access to digital data associated with the science, technology, and innovation enterprise. Historically, most of the studies from which such data are derived have been econometric or "scientometric" in nature, focusing on the development of quantitative data, models, and metrics of the scientific process as well as outputs and outcomes. Broader definitions of research impact, however, necessitate the use of qualitative case-study methods. For many years, U.S. federal science agencies such as the National Institutes of Health have demonstrated the impact of the research they support through tracing studies that document critical events in the development of successful technologies. A significant disadvantage and barrier of such studies is the labor-intensive nature of a case study approach. Currently, however, the same data infrastructures that have been developed to support scientometrics may also facilitate historical tracing studies. In this paper, we describe one approach we used to discover long-term, downstream outcomes of research supported in the late 1970's and early 1980's by the National Institute of General Medical Sciences, a component of the National Institutes of Health.

Keywords: science of science, citation networks, knowledge diffusion, research evaluation, research outcomes, basic research, government funding

OPEN ACCESS

Edited by:
Rafael Repiso,
Universidad Internacional De La Rioja, Spain

Reviewed by:
Jonathan Grant,
King's College London, United Kingdom
Matthew Mayernik,
National Center for Atmospheric Research (UCAR), United States

****Correspondence:***
James Onken
james.onken@reaconsultant.com

Specialty section:
This article was submitted to Research Assessment, a section of the journal Frontiers in Research Metrics and Analytics

Received: *28 May 2020*
Accepted: *03 August 2020*
Published: *08 September 2020*

Citation:
Onken J, Miklos AC and Aragon R (2020) Tracing Long-Term Outcomes of Basic Research Using Citation Networks. Front. Res. Metr. Anal. 5:5. doi: 10.3389/frma.2020.00005

INTRODUCTION

For more than a decade, beginning when Dr. Jack H. Marburger III, the President's Science Advisor and Director of the Office of Science and Technology Policy called for a "science of science (SoS) policy" (Office of Science Technology Policy, 2006), there has been a growing community of practice in the US surrounding the evaluation of scientific research programs. Interest in SoS was both reflected in, and further stimulated by, creation of the Science of Science Innovation and Policy (SciSIP) program at the National Science Foundation (NSF) in 2006 (National Research Council, 2014a). Through its grant program, SciSIP fostered the development of data, tools, and methods "to inform the nation's public and private sectors about the processes through which investments in science and engineering (S&E) research are transformed into social and economic outcomes" (National Science Foundation, 2007). The importance of these activities is further strengthened by the involvement of other federal agencies in SciSIP, such as the National Institute of General Medical Sciences (NIGMS), a component of the National Institutes of Health (NIH) (National Institute of General Medical Sciences, 2019a). Interest in SoS—also sometimes referred to as "research on research" or "meta-research" (Kamenetzky and Hinrichs-Krapels, 2020)—has not been limited to the US. Twelve countries and regions from around the world are partners in the Research on Research Institute, established in 2019 by the Wellcome Trust, Digital Science, and the Universities of Sheffield and Leiden (Skelton, 2019).

The increase in SoS studies has been fueled, in part, by greater access to digital data on the science, technology, and innovation enterprise (National Research Council, 2014b; Fortunato et al., 2018; Waldman and Lariviere, 2020). As more sophisticated databases, tools, and methods have become available, expectations—and sometimes requirements—for public science funding agencies to document the outcomes of national investments in research have increased (Husbands Fealing et al., 2011; Oancea, 2013; Kamenetzky and Hinrichs-Krapels, 2020).

Some agencies have responded by strengthening their own data infrastructure to facilitate SoS studies. In the UK, routine collection of research impact data has expanded through the use of national databases such as researchfish® (Raftery et al., 2016). In the US, the NIH has been leading the effort of several science agencies to construct the Federal Research Portfolio Online Reporting Tools: Expenditures and Results (RePORTER) website, a database of federal research investments and associated outputs (scientific publications) (National Institutes of Health, 2019). Federal RePORTER was modeled on NIH's own RePORTER system, which links NIH-funded projects to resulting publications and patents (National Institutes of Health, 2020a). Also, the NIH Office of Portfolio Analysis has created both internal and publicly available portfolio analysis tools and data, such as the NIH Open Citation Collection (Hutchins et al., 2019) and iCite, a query and analysis tool (National Institutes of Health, 2020b). The NIH Office of Extramural Research also has created an internal NIH Portfolio Analysis and Reporting Data Infrastructure (PARDI) that combines grant records, NIH-supported publications and patents, and citation data for use by NIH staff (Zuckerman et al., 2015).

Historically, SciSIP has been largely focused on econometric or "scientometric" research: the development of quantitative data, models, and metrics of the scientific process, outputs, and outcomes (National Academies of Sciences, 2017). There have been long-standing concerns surrounding the interpretation and use of some metrics (Donovan, 2007), and a rise in their application coincided with the creation in 2015 of the Leiden Manifesto, a set of principles to guide the use of metrics so that "researchers can hold evaluators to account, and evaluators can hold their indicators to account" (Hicks et al., 2015). Despite the SoS community's increased focus on metrics, the first principle in the Manifesto emphasizes the primacy of qualitative assessment, which quantification can support but not replace.

Broader definitions of research "impact" beyond economic measures to include social, cultural, and environmental returns have also necessitated the use of qualitative case-study methods (Kearnes and Wienroth, 2011), such as the Payback Framework, which has been used in several countries to assess the impact of health-related research (Buxton and Hanney, 1996; Donovan, 2011; Donovan and Hanney, 2011). Case studies formed the basis for the UK's Research Excellence Framework beginning in 2014 (King's College London Digital Science, 2015; Research Excellence Framework, 2015). That same year in the US, the NIH Scientific Management Review Board, charged with reviewing approaches to assess the value of biomedical research, concluded that "[n]arratives constructed from well-designed case studies can be especially effective illustrations of the broad impacts of biomedical research" (National Institutes of Health, 2014). In a similar vein, NSF recently changed the name and focus of the SciSIP program to "Science of Science: Discovery, Communication, and Impact," which may signal less emphasis being placed on metrics and an increase in the program's focus on how to enhance the value of scientific research to the public and stakeholders (National Science Foundation, 2019).

Case studies have long been used by public science funding agencies to demonstrate the impact of the research they support. One approach commonly used is the "historical tracing" or "historiographic" method (Ruegg and Jordan, 2007), a narrative account of the value of research in creating downstream inventions, products, or social benefits by tracing a series of incremental scientific advances ending in some outcome of value, such as improved public health. Tracing studies have a long history. In the late 1960's, the US NSF supported the TRACES (Technology in Retrospect and Critical Events in Science) study, which illustrated the role of basic research in five significant technologies, including the video tape recorder, oral contraceptives, and the electron microscope (Narin, 2013). The TRACES study was a response to "Project Hindsight" a similar study conducted by the US Department of Defense to assess the impact of its basic research (Sherwin and Isenson, 1967). More recent examples of tracing studies include those produced in the US by the Centers for Disease Control and Prevention (Centers for Disease Control Prevention, 2017) and the NIH (National Institutes of Health, 2018a).

A significant disadvantage of tracing studies, and a barrier to their use among science agencies, is the labor-intensive nature of the method (Comroe and Dripps, 1976; Smith, 1987; Contopoulos-Ioannidis et al., 2003; Mayernik et al., 2016). Expert knowledge is critical in identifying significant events in the path from basic research to societal outcomes (Narin, 2013; Centers for Disease Control Prevention, 2017). Such expertise can be costly, whether it is in terms of federal staff time or the cost of hiring expert consultants. However, the same data infrastructures that have been developed to support scientometrics may also facilitate historical tracing studies. The manual search for, and documentation of, evidence that basic research has contributed to a significant scientific or technological advance might be facilitated by a data infrastructure consisting of linked databases having records of research grants, scientific publications, patents, and other artifacts captured throughout the research and development process.

A data infrastructure such as that described above might also help meet an even greater challenge: continuously monitoring downstream technological advances to understand whether or to what extent they might have drawn on the results of a specific portfolio of basic research. Many historical tracing studies begin with a significant advance and trace backwards to identify the research on which it was based. For example, to demonstrate the impact of its research, NIH began with the development of childhood Haemophilus influenzae type b vaccines and worked backwards to identify prior vaccine development and the foundational research supported by NIH (National Institutes of Health, 2018b). Even when linked data sources are used, the tracing process has typically begun with the endpoint and worked backwards [see for example, Williams et al. (2015), Keserci et al. (2017)].

In contrast, forward tracing is a process of discovery beginning with a well-defined set of inputs whose outcomes have yet to be identified [see, for example, Wooding et al. (2011)]. A portfolio of research, embodied in a group of research grants or journal articles, can be traced through multiple generations of references to that work in subsequent journal articles, patents, clinical trials, clinical practice guidelines, drug products, etc.

One challenge in forward tracing is the exponential nature of knowledge diffusion (Chen and Hicks, 2004). Even a small number of research projects or articles, traced over a long period of time, can create a large amount of data that must be analyzed to identify significant outcomes. For example, in one study, an initial cohort of only 29 papers was cited by 731 unique second-generation papers ("unique" meaning second-generation papers that were not in the initial cohort), which were cited by 9,376 unique papers in the third generation (Hanney et al., 2005). There are currently no standard procedures or best practices to perform the data reduction and other processing necessary to identify significant outcomes or intermediates of interest that might be found among the large base of knowledge flowing from a particular portfolio of research. In this respect, the current state of the art is analogous to the ever-increasing volume of

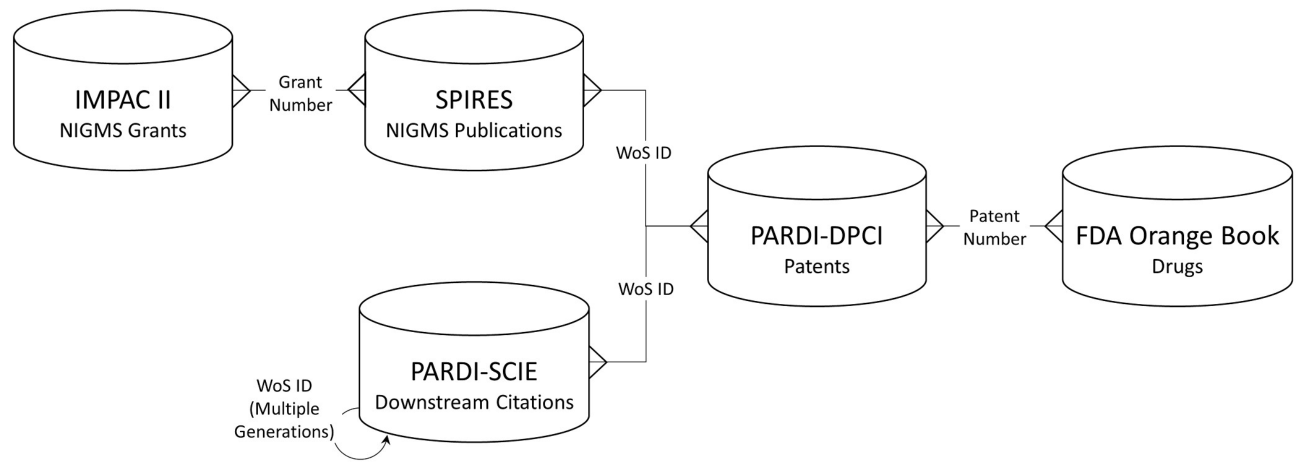

FIGURE 1 | Databases used and linkage keys. IMPAC II, NIH Information for Management Planning Analysis and Coordination; SPIRES, NIH Scientific Publication Information Retrieval and Evaluation System; PARDI-SCIE, Clarivate Science Citation Index Expanded, included in the NIH Portfolio Analysis and Reporting Data Infrastructure; PARDI-DPCI, Clarivate Derwent Patent Citation Index, included in PARDI; WoS, Web of Science.

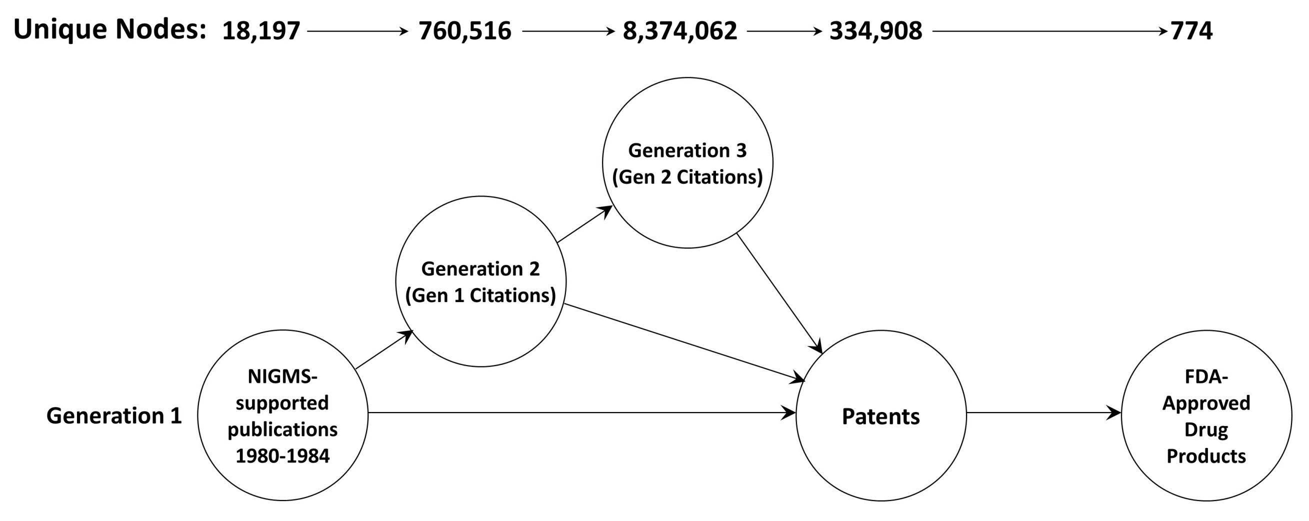

FIGURE 2 | The three-generation publication network. This is a forward trace of 18,197 publications in 1980–1984 citing support from NIGMS.

genomic sequence data, which has driven the need for enhanced bioinformatics tools necessary to analyze it (Batley and Edwards, 2009; Magi et al., 2010).

In this paper, we describe one approach we used to discover whether there are long-term, downstream technological advances to which research supported in the late 1970's and early 1980's by the NIGMS may have contributed. NIGMS administers a large portfolio of grants to support basic research in the biomedical sciences. In the five-year period from 1980 (the first year for which sufficient data are available) through 1984, over 18,000 publications cited support from NIGMS funding. We demonstrate one method by which significant health-related outcomes that are built on this research can be identified. In so doing, we also make some observations on the knowledge diffusion network created. This effort represents an initial attempt to define a replicable workflow that might be applied to other large portfolios of research and used routinely by other agencies and organizations to scan for significant outcomes as they occur.

MATERIALS AND METHODS

Data Sources

Long-term outcomes associated with NIGMS-funded research were identified through several types of linked data: NIGMS grants, publications citing NIGMS grant support, "downstream" publications that cited the NIGMS-supported publications, patents whose non-patent literature referenced either an NIGMS-supported or downstream publication, and drug products approved by the U.S. Food and Drug Administration (FDA) that are protected by one or more of the linked patents. **Figure 1** shows the data sources used and the structure of the network created among them.

Each of these data sources has certain weaknesses that could prevent a comprehensive and statistically accurate assessment of NIGMS-funded research outcomes. It is known that authors don't always acknowledge their grant support in the papers they publish (in the past, some journals have not permitted such acknowledgments) and the extent of such underreporting is not known. (The ability to link publications to NIH grant support has improved in recent years; in 2008, NIH began requiring reporting of grant-supported publications as a precondition for continued support.) Furthermore, when grant support is cited, it is prone to errors such as typographical mistakes in grant numbers. Similarly, references to non-patent literature in patents is prone to error, as these references are sometimes not detailed enough to uniquely identify the cited paper—for example, only an author and year of publication might be cited. Nor do we have complete information on patents associated with FDA-approved drug products. Of the 6,843 products named in the FDA Orange Book, patent information was available for only 16 percent. However, our primary goal in analyzing these datasets was not to generate a precise and reliable quantitative measurement of research outcomes, but rather to discover long-term outcomes that could be traced back to NIGMS-funded research, as the data allowed, and to enumerate any linkages found.

NIH Grants

Information on NIGMS grants was drawn from NIH's Information for Management Planning Analysis and Coordination (IMPAC) II database, an internal NIH database of grant applications and awards maintained by NIH's Office of Electronic Research Administration. While we used an internal database as our source data, a public version of the database is available (National Institutes of Health, 2017a, 2020a).

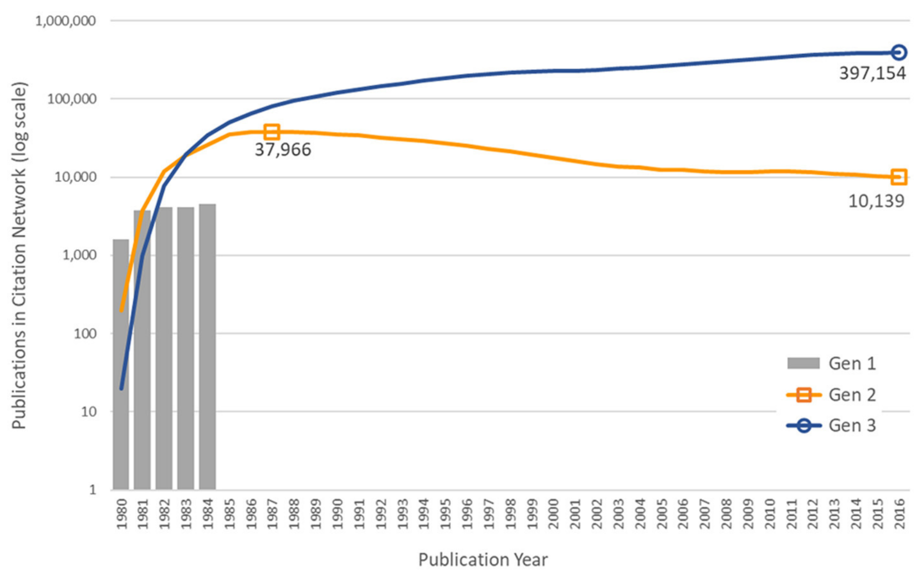

FIGURE 3 | Number of publications in the citation network, by generation and year of publication.

NIGMS Publications

The publications citing NIGMS grant support were retrieved from NIH's Scientific Publication Information Retrieval and Evaluation System (SPIRES). SPIRES is an internal database maintained by NIH's Office of Research Information Systems that relies on the Grant Support tag (GR) in MEDLINE/PubMed publication records (National Library of Medicine, 2019) to link publications to NIH grants. While we used this internal database for this study, a public version is available (National Institutes of Health, 2017b). Publications in SPIRES date from 1980. For this study, all publications from 1980–1984 citing support from NIGMS were selected as the starting point for the analysis.

Downstream Publications

Information on "downstream" publications—articles that have cited NIGMS publications—was obtained from NIH's Portfolio Analysis and Reporting Data Infrastructure (PARDI), a non-public NIH database that includes records from the Clarivate Analytics Science Citation Index Expanded® (SCIE). A recursive search of the SCIE can be performed to produce multiple generations of citations. All papers in the SCIE published from 1980 through 2016 were included in the analyses.

Patent Awards

Patents that include NIGMS and downstream publications in their non-patent literature references were also obtained from NIH's PARDI, which includes the Clarivate Analytics Derwent World Patents Index®.

Drug Products

Patent information on drug products was obtained from FDA's Approved Drug Products with Therapeutic Equivalence Evaluations (Orange Book) Data Files (U.S. Food Drug Administration, 2020). First published in 1980, the "Orange Book" identifies all currently marketed drug products approved on the basis of safety and effectiveness by the FDA. The February 2019 version of the Orange Book was used in the analyses. At that time, there was a total of 6,843 drug products with distinct trade names in the Orange Book. Patent information was available for 1,079 of these products.

RESULTS

The Knowledge Diffusion Network

Previous research has suggested the diffusion of knowledge underlying scientific progress is captured best by multiple

TABLE 1 | Number of publications in each generation cited by patents.

Generation	Publications	Cited by patents	% Cited	Citing patents	Patent-pub pairs	Avg citations
1	18,197	3,199	17.58	16,452	26,030	8.14
2	760,516	127,132	16.72	159,378	869,666	6.84
3	8,374,062	673,977	8.05	319,481	3,473,972	5.15
Total	9,152,775	804,308	8.79	334,908	4,369,668	5.43

Average citations calculated using only publications cited by at least one patent.

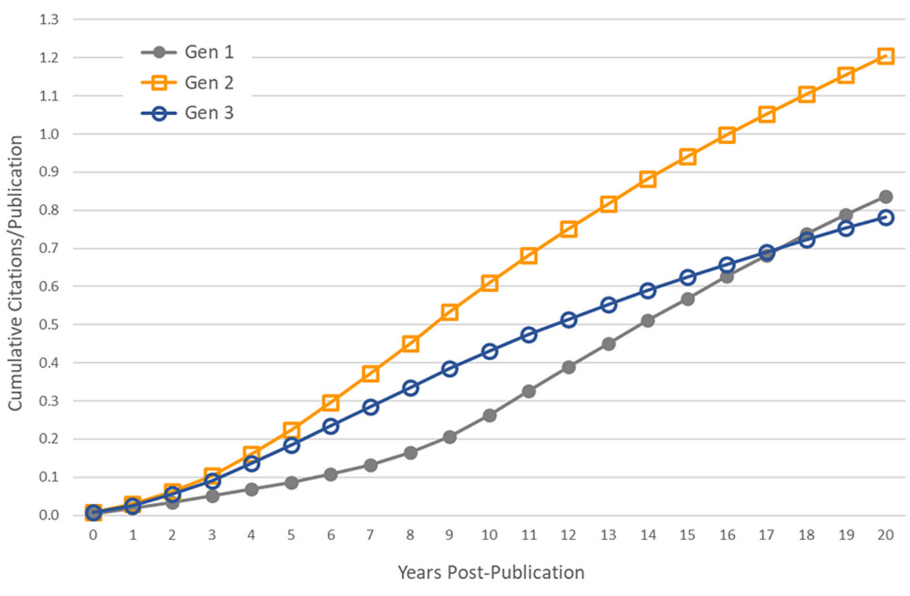

FIGURE 4 | Controlling for censored time series data. Average cumulative number of patent citations to papers published in the year 1993 and earlier, by generation.

generations of citations (Hu et al., 2011). However, we found little in extant literature to guide our choice of how many generations to include in our analysis. The degree and type of impact properly attributable to research when its influence is exerted indirectly through multiple generations of citations is not clear. There are several characteristics that could affect the number of generations that should be included to assess impact and the need for more research on this topic has been noted (Fragkiadaki and Evangelidis, 2016). We found examples of previous research using four to six generations of publications to trace the long-term impact of biomedical research (Grant et al., 2000, 2003; Jones and Hanney, 2016).

In this study we traced the first generation of NIGMS-funded articles forward for two subsequent generations of literature citations to find links to patents. We expected few articles in the first generation of NIGMS-supported publications to be directly cited by patents; the mission of NIGMS is to support research into fundamental biological processes. The Institute does not fund research directly related to a specific disease, life stage, population, or organ system—research which is supported by the other "categorical" NIH institutes and centers (National Institute of General Medical Sciences, 2019b). We considered it more likely that the role of NIGMS research in patented inventions would be found through later generations of research articles that built upon and cited NIGMS-funded research. However, as more generations of publications are added to the network, the relevance of the original NIGMS-funded research to any patent citing that literature may become more tangential. To focus on those patents to which NIGMS research may have contributed most directly, we limited the citation network to only three generations. In previous research, three generations have been considered sufficient to illustrate the usability and feasibility of various measures of impact (Fragkiadaki and Evangelidis, 2016).

A total of 18,197 articles published in 1980–1984 cited support from NIGMS (**Figure 2**). The second generation consisted of 760,516 unique papers and a third generation of 8,374,062 articles cited one or more of the second-generation papers. A total of 334,908 different patents cited at least one article from these three generations of papers. There were 774 different drug products that claimed protection from these patents, representing 11.3 percent of the 6,843 unique trade-named drug products and 71.8 percent of the 1,078 products having patent information in the Orange Book.

Citing Publications

The numbers of publications by year and generation number are shown in **Figure 3**. As discussed above, the first generation consists of 18,197 articles published in 1980–1984 citing NIGMS support. The number of articles citing the first generation each year reached a peak of 37,966 in 1987, an average of 4.65 years after the NIGMS papers were published. In 2016, the first-generation NIGMS papers were still being cited over 10,000 times. The third-generation papers had not peaked by 2016, when there were 397,154 articles citing one or more the 760,516 papers in the second generation.

Citing Patents

Table 1 shows the number of publications in each generation that were cited in patents' non-patent literature. Of the 1980–1984 publications citing NIGMS support, 17.58 percent were cited by at least one of 16,452 patents. The cited publications were referenced in an average of about eight patents.

Subsequent generations of publications were less likely to be cited by a patent, and those papers that were cited were referenced on fewer patents. However, these statistics are influenced by the censored distributions of the second and third generations of articles. Many of these articles have been published in recent years and some will be cited by patents in the future. To control for this effect, we used only papers published in the year 1993 and earlier—providing a citation follow-up time of at least 20 years for all papers—and calculated the average cumulative number of patent citations that papers received in the first 20 years post-publication. These cumulative distribution functions are shown in **Figure 4**. In general, the first generation of papers, which we expect to be more heavily weighted toward basic research, have fewer patent citations in the years immediately following publication, but they are cited at a higher rate over longer periods of time than second- and third-generation papers, eventually surpassing generation 3.

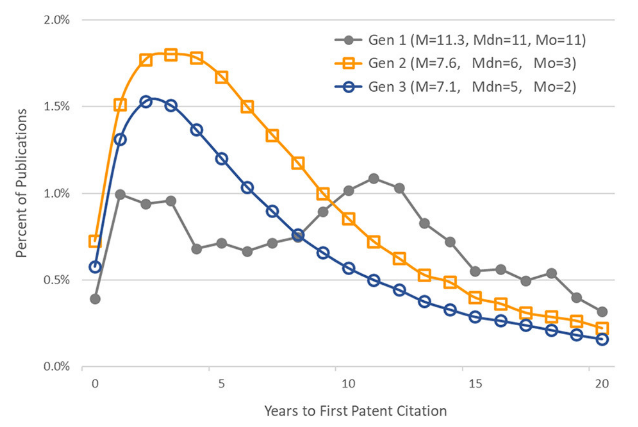

FIGURE 5 | Distribution of publications by time to first patent citation. Generation 1 articles published in 1980–1984. Generations 2 and 3 articles published in 1980–1993. Also noted are the mean (M), median (Mdn), and mode (Mo) of each generation's distribution.

TABLE 2 | U.S. sales in 2018 for top-selling drugs that were linked to NIGMS-supported research.

Rank	Trade_name	Publication generation	Sales ($B)
8	Imbruvica	2	4.10
12	Genvoya	3	3.63
13	Lyrica	3	3.59
16	Ibrance	3	2.90
19	Victoza	2	2.70
20	Truvada	2	2.60

Time to First Patent Citation

Figure 5 shows, for each generation, the distribution of publications by number of years to first patent citation. The mean (M), median (Mdn), and mode (Mo) are also given for each generation's distribution.

Linked Products

Of the 774 products linked to patents in the NIGMS citation network, six (shown in **Table 2**), were among the top 20 best-selling drugs in the U.S. for 2018 (Questex, 2020). The total sales of these six products was $19.52 billion. "Publication Generation" indicates the generation in which these products' patents entered the citation network.

Data Reduction

This high-level summary of the citation network provided us with descriptive information on the broad diffusion of knowledge developed through NIGMS-funded research. However, our primary interest is in how to distill this large amount of information to identify *specific* outcomes of interest and significant events in the research and development process. We turned to the network analysis and visualization platform Gephi (Bastian et al., 2009) to analyze the network and locate nodes of significance.

Unfortunately, a network of the size we originally created, with 9.5 million nodes, exceeds the capacity of Gephi (as well as some other popular graph visualization and analysis tools; Pavlopoulos et al., 2017). As a result, we included only two generations of publications—the original set of 18,197 articles supported by NIGMS and all the papers that cited one or more of these NIGMS publications—and, of these, only publications eventually led to a patent. The numbers of nodes in this reduced network and edges are shown in **Figure 6**. The NIGMS publications were ultimately linked to 435 drug products,

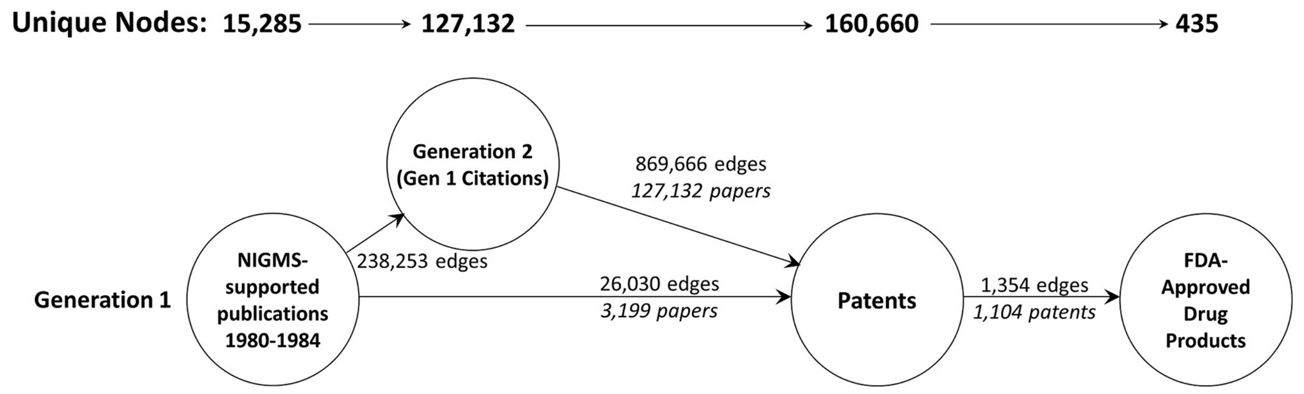

FIGURE 6 | The two-generation publication network. A forward trace of 15,285 publications in 1980–1984 citing support from NIGMS that eventually led to one or more patents.

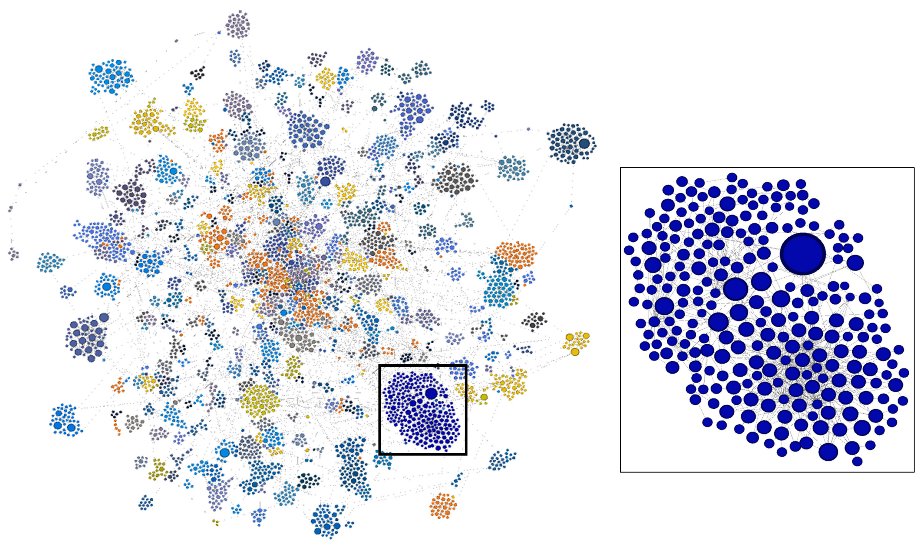

FIGURE 7 | Complete two-generation network as a directed graph. Highly linked clusters are identified by color. Blue cluster in box is enlarged at right. Visualization by Gephi.

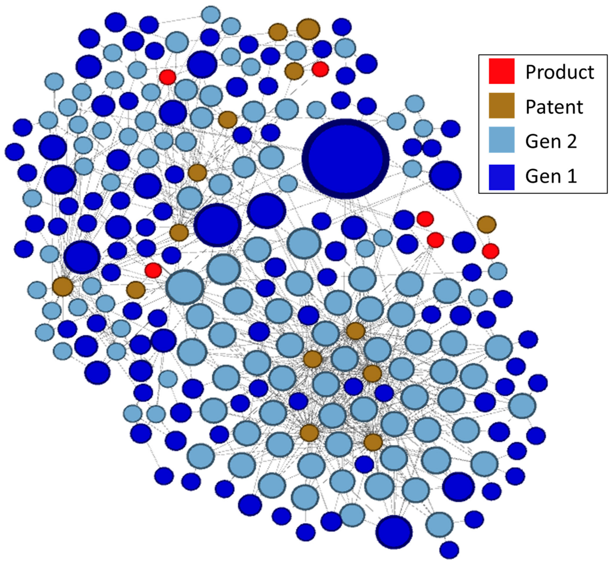

FIGURE 8 | Highly linked cluster shown in **Figure 7**, by type of node.

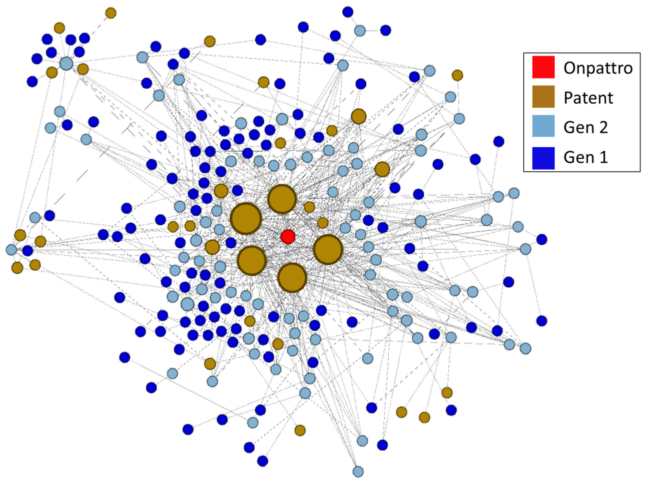

FIGURE 9 | Patents and publications in the Onpattro cluster. Seven patents are associated with Onpattro in the FDA Orange book. The cluster also contains seven additional patents not related to Onpattro.

representing 6.4 percent of the 6,843 unique trade-named drug products and 40.4 percent of the 1,078 products having patent information in the Orange Book.

A visualization of this network as a directed graph is shown in **Figure 7**. To discern meaningful relationships or patterns in the network, we identified clusters of related research, patents, and products using the Louvain Method of community detection (Blondel et al., 2008), implemented in Gephi using a randomized parameter, no weights, and 1.0 resolution. Network visualization was performed with Gephi 0.9.1 using the ForceAtlas2 layout algorithm and default parameters.

In the lower right portion of the network in **Figure 7** is a cluster that appears to be particularly large and whose nodes are highly linked. We arbitrarily selected this cluster for further analysis. In **Figure 8**, this cluster is isolated and color-coded by node type, making it easier to identify patents and drug products that are linked to the outputs of research supported by NIGMS. The size of the blue publication nodes is proportional to the number of times each has been cited by other publications or patents (indegree). The size of the brown patent nodes is also proportional to their indegree, the number of drug products for which they provide intellectual property protection.

Developing a Product Trace

There were 14 patents in the cluster, related to six drug products, all of which affect gene expression, including several oligonucleotide therapeutics, a relatively new class of drugs made of chemically synthesized nucleic acids (Smith and Zain, 2019):

> Cubicin® RF (daptomycin) is a last-resort antibiotic with excellent activity against Gram-positive pathogens. It was first approved by the FDA for use in the treatment of skin infections. It has a distinct mechanism of action causing rapid depolarization of membrane potential, disrupting cell membrane function to inhibit protein, DNA, and RNA synthesis;

> Epiduo® (adapalene and benzoyl peroxide) is a treatment for severe acne. Adapalene binds to retinoic acid nuclear receptors, which act as transcription factors to regulate the expression of mRNA for proteins modulating cell differentiation and keratinization;

> Kynamro® (mipomersen) is an adjunct to lipid-lowering medications to reduce LDL in patients with homozygous familial hypercholesterolemia. It is an antisense oligonucleotide targeted to human messenger ribonucleic acid (mRNA) for apo B-100, the principal apolipoprotein of LDL;

> Onpattro® (patisiran) is a small interfering RNA (siRNA) oligonucleotide for the treatment of polyneuropathy in people with hereditary transthyretin-mediated amyloidosis. It is the first siRNA-based drug approved by the FDA;

> Tegsedi™ (inotersen) is for the treatment of polyneuropathy in people with hereditary transthyretin-mediated amyloidosis. It also is an antisense oligonucleotide that inhibits hepatic production of transthyretin by binding to mRNA;

> Zemdri™ (plazomicin) is an aminoglycoside antibacterial for the treatment of complicated urinary tract infections. It acts by binding to bacterial 30S ribosomal subunits, interfering with mRNA and protein synthesis.

The linkage of NIGMS-funded research to Onpattro, being the first siRNA-based drug and only recently approved for use, was a particularly interesting discovery. We reduced the data further by examining the nodes in the immediate neighborhood of Onpattro. This smaller network is shown in **Figure 9**, where the nodes have been resized according to their indegree within the Onpattro network. The single product node, Onpattro, is colored red and located in the center. It is surrounded by seven patents, in brown, linked to NIGMS-funded research that are associated with Onpattro in the FDA Orange Book (out of a total of 21 patents for Onpattro). These seven patents fall into three families (**Table 3**). In the Government Interest section of one of these patent families, *RNA sequence-specific mediators of RNA interference*, support is acknowledged from NIGMS grant number GM034277, a grant on the regulation of mRNA processing awarded to Philip Sharp, a Nobel Laureate who co-discovered RNA splicing.

The five patents with the largest indegree (i.e., largest number of connections to the NIGMS publication network), evident in **Figure 9**, are from two families: *RNA sequence-specific mediators of RNA interference* and *RNA interference mediating small RNA molecules*. The periphery of **Figure 9** also shows seven patents citing literature in the Onpattro network but which are not themselves linked to Onpattro in the Orange Book.

To establish a chronology of events involved in the development of Onpattro, the network nodes were placed on a timeline using publication years, patent application and award dates, and the approval of Onpattro in 2018 (**Figure 10**). The timeline includes the network's 67 first-generation publications in 1980–1984 (median publication year = 1983), 62 second-generation publications (median publication year = 1998), and the seven patent applications (median date = 2010) and awards (median date = 2014). Of the 62 second-generation publications, 19 resulted from NIGMS-funded research. Across both generations, NIGMS supported 86 (66.7 percent) of the 129 publications in the Onpattro network.

NIGMS Grant Support Linked to Onpattro Development

Finally, we identified the specific NIGMS-funded basic research that produced many of the articles in the Onpattro network. A total of 80 NIGMS grants were cited by publications in the Onpattro network. Ten of these were responsible for generating 32 first- and second-generation articles (**Table 4**) representing 37 percent of all NIGMS-supported publications and almost one-quarter of all articles in the network. Eight of these grants generated 34.3 percent of the first-generation publications from which the knowledge dissemination network was developed. Two of the ten grants were awarded after 1984 and supported only publications in the second generation. All of these grants involve well-known investigators, including three Nobel Laureates, working in areas of basic research critical to the development of oligonucleotide therapeutics and other drugs affecting gene expression. Also shown in **Table 4** is the total amount of NIGMS funding for these grants through 1984, the final year of the first-generation publications used to generate the Onpattro network. Total NIGMS support for the first-generation publications was $206 million.

DISCUSSION

This initial attempt to interrogate a large network of documentary evidence, beginning with the results of basic research funded by NIGMS in 1980–1984, provided us with several interesting findings to be explored in more depth. The knowledge flowing from this body of research was traced to the development of 774 drug products, including some of the most popular drugs in use today. Six of these drugs were among the best-selling in 2018, with sales of $19.5 billion (Questex, 2020). In comparison, the total funding for the 48 research grants producing this research, through 1984, was $82.9 million, ∼311.7 million in 2018 dollars (National Institutes of Health, 2020c). Limiting grant funding to the five years preceding publication—research in the year of publication and the four prior years—reduces the total NIGMS investment that gave rise to this knowledge diffusion network to $44 million, ∼$165 million in 2018 dollars.

We lack sufficient data to calculate a return on investment (ROI) from these figures. A proper calculation of ROI would require more completely identifying all non-NIGMS inputs contributing to long-term outcomes, including negative outcomes (i.e., revenue losses), applying an appropriate economic valuation to the outcomes, and weighing the attribution of outcomes to each input (Buxton et al., 2004). Previous studies have more rigorously produced estimates of the economic returns of funding for health-related research (Buxton et al., 2004; RAND Europe., 2008; Grant and Buxton, 2018). We will simply note that, while NIGMS funding for basic research is only one portion of the total required to bring these products to market, these fundamental discoveries were critical to drug development and the amount of NIGMS funding required was small relative to the value of the outcomes to which they contributed.

We were also able to gain a better sense of the amount of time required for the diffusion of the basic research findings generated by NIGMS funding. Second-generation citations of research published in 1980–1984 peaked in 1987, an average of 4.65 years after publication [consistent with previous findings, see Fukuzawa and Ida (2016)], but there were still many citations of this work in 2016, more than thirty years later. It was also interesting to see that third-generation citations of this research had not peaked by 2016 and continue to grow in number.

Carpenter et al. (1980) found that patents cite relatively recent literature, but these citations varied by technology area; the median "age" of articles cited (the time elapsed from the paper's publication to its patent citation) by gas laser patents was only three years. It was slightly longer in prostaglandin patents. Chen and Hicks (2004), studying tissue engineering research articles, found the time elapsed between papers' publication year and their first front-page citation in a patent had a mean of 9.6 years, and a mode of 2 years. While patents may more frequently cite recent research, it is not necessarily the case that most of a research article's patent citations will occur shortly after publication.

In our study, the time elapsed from an article's publication and its first citation in a patent application varied as a function of a publication's generation in the citation network. We provided a minimum of 20 years follow-up for articles in all generations. As expected, the time elapsed from the original basic research funded by NIGMS to its first citation in a patent was longer than for later generations of publications that built on this research. The median time to first patent citation in the first

TABLE 3 | Onpattro patent families and number found in the NIGMS network.

Patent family	Patent year(s)	Patents in family	Patents in NIGMS network	Government interest
Lipid formulations for nucleic acid delivery	2011–2016	4	0	
Lipid formulation	2012, 2014	2	0	
Nuclease resistant double-stranded ribonucleic acid	2012	1	0	
Compositions and methods for inhibiting expression of transthyretin	2012–2016	3	2	
RNA sequence-specific mediators of RNA interference	2013, 2015	2	2	GM034277
RNA interference mediating small RNA molecules	2013–2017	6	3	
Lipid containing formulations	2014	1	0	
2′-methoxy substituted oligomeric compounds and compositions for use in gene modulations	2018	2	0	

Onpattro patents from U.S. Food Drug Administration (2020).

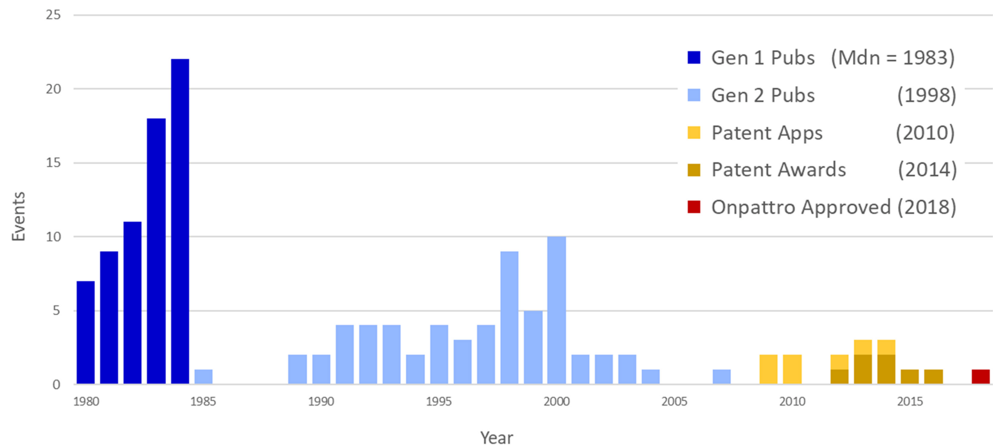

FIGURE 10 | Timeline of events in the Onpattro development network.

TABLE 4 | NIGMS research grants generating 37 percent of all NIGMS-supported publications associated with Onpattro development.

Investigator	Gen1 pubs	Gen2 pubs	Grant	Grant title	First funded	Funding through 1984
Fire, Andrew	0	7	GM037706	Gene regulation during early development of C elegans	1987	N/A
Horvitz, H Robert	5	1	GM024663	Genetic analysis of nematode egg-laying	1978	$28,497,885
			GM024943	control of cell division in the nematode C elegans	1978	$29,721,139
Apirion, David	5	0	GM019821	Genetics and biochemistry of RNA processing in *E. coli*	1974	$27,066,386
			GM025890	The molecular biology of RNA turnover in *E. coli*	1979	$10,185,015
Hershey, John	4	0	GM022135	Mechanism of initiation of protein biosynthesis	1975	$30,604,835
Levin, Daniel	3	0	GM024825	Control of protein synthesis by double-stranded RNA	1978	$29,321,169
T'so, Paul	3	0	GM016066	Nucleic acid chemistry and its biomedical application	1970	$19,557,228
Turner, Douglas	3	0	GM022939	Kinetic and spectroscopic studies of nucleic acids	1976	$30,598,690
Sharp, Philip	0	1	GM034277	Regulation of MRNA processing	1985	N/A
Total	23	9				$205,552,347

generation of 1980–1984 articles was 11 years. The lags for the second- and third-generation papers were shorter, at 6 and 5 years, respectively. The lag distributions for generations 2 and 3 were similar to those found by Chen and Hicks (2004).

Using clustering techniques (which required that we use a smaller dataset excluding third-generation citations), we were able to discover meaningful clusters of papers, patents, and products. One cluster, selected arbitrarily, was a network of documents related to a class of drugs affecting gene expression, including several recently-developed oligonucleotide therapeutics. We were able to identify the specific NIGMS research grants and investigators whose research contributed to the development of one of these drugs, Onpattro. From the network of linked documents, we were able to chart key events occurring over a 38-year period from the first NIGMS-supported publications in 1980-1984 to approval of Onpattro in 2018. About one-half of the NIGMS-supported articles were produced by grants totaling about $206 million through 1984. It is not clear how much of this funding would have directly contributed to these key publications; for example, some of these grants began in the early 1970s. By contrast, net product revenues for Onpattro through 2019 were $179 million and sales of $285 to $315 million are forecast for 2020 (Alnylam, 2020).

The Onpattro example demonstrates the ability to easily discover useful new knowledge from large linked datasets of information. The automated procedures to do so may provide a useful alternative to the labor-intensive approaches that have been used in the past, such as that described in Comroe and Dripps (1976). While the validity of their findings has been challenged (Smith, 1987; Grant et al., 2003), we cite Comroe and Dripps only to exemplify the effort that can be required by this type of study. The Comroe and Dripps study employed about 200 consultants to identify the most significant advances in cardiovascular and pulmonary medicine, the essential bodies of knowledge required for these advances, and key articles in these knowledge flows. They used the expert opinion of many consultants to avoid bias in the selection of key prior research and articles. More recent tracing studies have continued to rely on labor-intensive methods (Smith, 1987; Contopoulos-Ioannidis et al., 2003; Mayernik et al., 2016). We employ a different approach to avoiding bias by using objective linkages among documentary evidence, and our study also employed a single primary researcher with access to the necessary databases and analytic tools.

Just as Chen and Hicks' work was the start of a long-term program to develop new analytic methods capturing knowledge diffusion (Chen and Hicks, 2004), we view our study as an exploratory effort to assess the utility of linked databases in tracing the long-term influence of a program of basic research in the biomedical sciences. Our ability to link new technologies to NIGMS-funded research was facilitated by an existing data infrastructure (PubMed) that links publications to NIH grants in the biomedical sciences (National Library of Medicine, 2019). Such resources are beginning to be made available in more areas of science. Other agencies, funding research in diverse fields, have started to make publications and other research products associated with their grants available through public data sources such as Federal RePORTER, an online searchable database developed by STAR METRICS, a consortium of US science agencies (Onken, 2016). Beginning in 2019, the EU has made project-linked publications available through its Open Data Portal (European Commission, 2020) and the UK provides data on publications linked to projects funding by nine agencies funding research in a range of fields (UK Research Innovation, 2020). Commercial bibliographic data services such as Web of Science, SCOPUS, and Dimensions have been capturing such information from funding acknowledgments in papers for some time (Rigby, 2011; Hook et al., 2018). Even in the absence of project identifiers linked to publications, our procedures allow the search for long-term outcomes to begin with any set of publications produced by a portfolio of research.

Procedures like those we describe here offer objective, reliable, and less time-consuming ways to discover knowledge flows contributing to new technologies and the research playing a critical role (van Raan, 2017). Our approach is, however, only one of many possible approaches. More research is needed to find other, more optimal approaches for linking databases, identifying critical nodes in knowledge flows, and exploring the meaningfulness of the networks discovered. Greater understanding is needed of the degree and type of impact properly attributable to research when its influence is exerted indirectly through multiple generations of citations. This will require more in-depth study that builds on the initial effort presented in this paper. For example, previous research has demonstrated a shift from basic to clinical science across forward generations of citations (Grant et al., 2000, 2003). If corroborated using the citation network developed in our study, we might be able to describe with greater specificity the contributions made by basic research. By investing in such research, automated procedures thus developed can be quickly and easily applied to other research programs, significantly reducing the time and effort required to demonstrate, in an objective way, long-term contributions of the results flowing from basic research programs.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. The first generation of NIGMS-supported publications and grant numbers, which formed the foundation of the citation network created in this paper, is publicly available at https://doi.org/10.6084/m9.figshare.12671045. Subsequent generations of publications in the citation network were obtained from the proprietary Clarivate Web of Science database and our license restricts distribution. The FDA Orange Book data used in this study can be found here: https://www.fda.gov/drugs/drug-approvals-and-databases/orange-book-data-files.

ETHICS STATEMENT

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

JO designed the study, performed data analysis, and wrote the initial draft manuscript. AM assisted in conceptualizing the study and revised the manuscript critically. RA secured funding, oversaw data interpretation, and revised the manuscript critically. All authors read and approved the final manuscript.

FUNDING

This work was supported in part by contract HHSN316201300006W/HHSN26300001 from the National Institute of General Medical Sciences to Medical Science & Computing, LLC. This support was used to procure the services of JO.

ACKNOWLEDGMENTS

We thank George Chacko and Dmitriy Korobskiy (NET Esolutions Corporation) for their assistance in working with the PARDI database.

REFERENCES

Alnylam (2020). *Alnylam Pharmaceuticals Reports Fourth Quarter and Full Year 2019 Financial Results and Highlights Recent Period Activity*. Available online at: https://investors.alnylam.com/press-release?id=24491 (accessed February 19, 2020).

Bastian, M., Heymann, S., and Jacomy, M. (2009). "Gephi: an open source software for exploring and manipulating networks," in *Third International AAAI Conference on Weblogs and Social Media* (San Jose, CA), 361–362.

Batley, J., and Edwards, D. (2009). Genome sequence data: management, storage, and visualization. *Biotechniques* 46, 333–336. doi: 10.2144/000113134

Blondel, V. D., Guillaume, J-L, Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp*. 10, 10008–10019. doi: 10.1088/1742-5468/2008/10/P10008

Buxton, M., and Hanney, S. (1996). How can payback from health services research be assessed? *J. Health Serv. Res. Policy* 1, 35–43. doi: 10.1177/135581969600100107

Buxton, M., Hanney, S., and Jones, T. (2004). Estimating the economic value to societies of the impact of health research: a critical review. *Bull. World Health Organ* 82, 733–739.

Carpenter, M. P., Cooper, M., and Narin, F. (1980). Linkage between basic research literature and patents. *Res. Manage*. 23, 30–35. doi: 10.1080/00345334.1980.11756595

Centers for Disease Control and Prevention (2017). *About the Science Impact Project*. Available online at: https://www.cdc.gov/od/science/impact/about.html (accessed February 3, 2020).

Chen, C., and Hicks, D. (2004). Tracing knowledge diffusion. *Scientometrics* 59, 199–211. doi: 10.1023/B:SCIE.0000018528.59913.48

Comroe, J. H., and Dripps, R. D. (1976). Scientific basis for the support of biomedical science. *Science* 192, 105–111. doi: 10.1126/science.769161

Contopoulos-Ioannidis, D. G., Ntzani, E. E., and Ioannidis, J. P. A. (2003). Translation of highly promising basic science research into clinical applications. *Am. J. Med*. 114, 477–484. doi: 10.1016/S0002-9343(03)00013-5

Donovan, C. (2007). The qualitative future of research evaluation. *Sci. Public Policy* 34, 585–597. doi: 10.3152/030234207X256538

Donovan, C. (2011). State of the art in assessing research impact: introduction to a special issue. *Res. Eval*. 20, 175–179. doi: 10.3152/095820211X13118583635918

Donovan, C., and Hanney, S. (2011). The "payback framework" explained. *Res. Eval*. 20, 181–183. doi: 10.3152/095820211X13118583635756

European Commission (2020). *EU Open Data Portal: Access to European Union Open Data*. Available online at: https://data.europa.eu/euodp/en/home (accessed July 8, 2020).

Fortunato, S., Bergstrom, C. T., Borner, K., Evans, J. A., Helbing, D., Milojevic, S., et al. (2018). Science of science. *Science* 359, 1007–1014. doi: 10.1126/science.aao0185

Fragkiadaki, E., and Evangelidis, G. (2016). Three novel indirect indicators for the assessment of papers and authors based on generations of citations. *Scientometrics* 106, 657–694. doi: 10.1007/s11192-015-1802-4

Fukuzawa, N., and Ida, T. (2016). Science linkages between scientific articles and patents for leading scientists in the life and medical sciences field: the case of Japan. *Scientometrics* 106, 629–644. doi: 10.1007/s11192-015-1795-z

Grant, J., and Buxton, M. J. (2018). Economic returns to medical research funding. *BMJ Open* 8:e022131. doi: 10.1136/bmjopen-2018-022131

Grant, J., Cottrell, R., Cluzeau, F., and Fawcett, G. (2000). Evaluating "payback" on biomedical research from papers cited in clinical guidelines: applied bibliometric study. *BMJ* 320, 1107–1111. doi: 10.1136/bmj.320.7242.1107

Grant, J., Green, L., and Mason, B. (2003). Basic research and health: a reassessment of the scientific basis for the support of biomedical science. *Res. Eval*. 12, 217–224. doi: 10.3152/147154403781776618

Hanney, S. R., Home, P. D., Frame, I., Grant, J., Green, P., and Buxton, M. J. (2005). Identifying the impact of diabetes research. *Diabetic Med*. 23, 176–184. doi: 10.1111/j.1464-5491.2005.01753.x

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., and Rafols, I. (2015). Bibliometrics: the leiden manifesto for research metrics. *Nature* 520, 429–431. doi: 10.1038/520429a

Hook, D. W., Porter S. J., and Herzog, C. (2018) Dimensions: building context for search and evaluation. *Front. Res. Metr. Anal*. 3:23. doi: 10.3389/frma.2018.00023

Hu, X., Rousseau, R., and Chen, J. (2011). On the definition of forward and backward citation generations. *J. Infometr*. 5, 27–36. doi: 10.1016/j.joi.2010.07.004

Husbands Fealing, K., Lane, J. I., Marburger, J. H., and Shipp, S. S. (2011). "Editor's introduction," in *The Science of Science Policy: A Handbook*, eds K. Husbands Fealing, J. Lane, J. Marburger III, and S. Shipp (Palo Alto, CA: Stanford University Press), 1–8.

Hutchins, B. I., Baker, K. L., Davis, M. T., Diwersy, M. A., Haque, E., Harriman, R. M., et al. (2019). The NIH open citation collection: a public access, broad coverage resource. *PLoS Biol*. 17:e3000385. doi: 10.1371/journal.pbio.3000385

Jones, T. H., and Hanney, S. (2016). Tracing the indirect societal impacts of biomedical research: development and piloting of a technique based on citations. *Scientometrics* 107, 975–1003. doi: 10.1007/s11192-016-1895-4

Kamenetzky, A., and Hinrichs-Krapels, S. (2020). How do organisations implement research impact assessment (RIA) principles and good practice? A narrative review and exploratory study of four international research funding and administrative organisations. *Health Res. Policy Syst*. 18:6. doi: 10.1186/s12961-019-0515-1

Kearnes, M., and Wienroth, M. (2011). Tools of the trade: UK research intermediaries and the politics of impacts. *Minerva* 49, 153–174. doi: 10.1007/s11024-011-9172-4

Keserci, S., Livingston, E., Wan, L., Pico, A. R., and Chacko, G. (2017). Research synergy and drug development: bright stars in neighboring constellations. *Heliyon* 3:e00442. doi: 10.1016/j.heliyon.2017.e00442

King's College London and Digital Science (2015). *The Nature, Scale and Beneficiaries of Research Impact: An Initial Analysis of Research Excellence Framework (REF) 2014 Impact Case Studies*. Bristol: HEFCE.

Magi, A., Benelli, M., Gozzini, A., Girolami, F., Torricelli, F., and Brandi, M. L. (2010). Bioinformatics for next generation sequencing data. *Genes* 1, 294–307. doi: 10.3390/genes1020294

Mayernik, M. S., Hart, D. L., Maull, K. E., and Weber, N. M. (2016). Assessing and tracing the outcomes and impact of research infrastructures. *J. Assoc. Inf. Sci. Tech*. 68, 1341–1359. doi: 10.1002/asi.23721

Narin, F. (2013). Tracing the paths from basic research to economic impact. *F&M Sci*. 2013, 67–88.

National Academies of Sciences, Engineering, and Medicine (2017). *Using Narrative and Data to Communicate the Value of Science: Proceedings of a Workshop—in Brief*. Washington, DC: The National Academies Press.

National Institute of General Medical Sciences (2019a). *NIGMS Feedback Loop: Apply to SCISIPBIO: A Joint Initiative Between NIGMS and NSF to Support Research on the Science of Science and Innovation Policy*. Available online at: https://loop.nigms.nih.gov/2019/02/apply-to-scisipbio-a-joint-initiative-between-nigms-and-nsf-to-support-research-on-the-science-of-science-and-innovation-policy (accessed April 9, 2020).

National Institute of General Medical Sciences (2019b). *About NIGMS: Overview*. Available online at: https://www.nigms.nih.gov/about/overview (accessed July 3, 2020).

National Institutes of Health (2014). *Scientific Management Review Board Report on Approaches to Assess the Value of Biomedical Research Supported by NIH*. Available online at: https://smrb.od.nih.gov/documents/reports/VOBR%20SMRB__Report_2014.pdf (accessed January 22, 2019).

National Institutes of Health (2017a). *Computer Retrieval of Information on Scientific Projects (CRISP) Legacy Data*. Available online at: https://exporter.nih.gov/crisp_catalog.aspx (accessed February 6, 2020).

National Institutes of Health (2017b). *ExPORTER Data Catalog: Publications*. Available online at: https://exporter.nih.gov/ExPORTER_Catalog.aspx?sid=0&index=2 (accessed February 6, 2020).

National Institutes of Health (2018a). *Impact of NIH Research: Our Stories*. Available online at: https://www.nih.gov/about-nih/what-we-do/impact-nih-research/our-stories (accessed January 29, 2020).

National Institutes of Health (2018b). *Childhood Hib Vaccines: Nearly Eliminating the Threat of Bacterial Meningitis*. Available online at: https://www.nih.gov/sites/default/files/about-nih/impact/childhood-hib-vaccines-case-study.pdf (accessed February 3, 2020).

National Institutes of Health (2019). *Federal Reporter*. Available online at: https://federalreporter.nih.gov (accessed January 29, 2020).

National Institutes of Health (2020a). *NIH Research Portfolio Online Reporting Tools: RePORTER*. Available online at: https://projectreporter.nih.gov (accessed January 29, 2020).

National Institutes of Health (2020c). *Biomedical Research and Development Price Index (BRDPI)*. Available online at: https://officeofbudget.od.nih.gov/gbipriceindexes.html (accessed March 11, 2020).

National Institutes of Health (2020b). *iCite*. Available online at: https://icite.od.nih.gov (accessed January 29, 2020).

National Library of Medicine (2019). *Funding Support (Grant) Information in MEDLINE/PubMed*. Available online at: https://www.nlm.nih.gov/bsd/funding_support.html (accessed February 6, 2020).

National Research Council (2014a). *Science of Science and Innovation Policy: Principal Investigators' Conference Summary*. Washington, DC: The National Academies Press.

National Research Council (2014b). *Capturing Change in Science, Technology, and Innovation: Improving Indicators to Inform Policy*. Washington, DC: The National Academies Press.

National Science Foundation (2007). *Science of Science Innovation and Policy FY 2007 Program Solicitation NSF 07-547*. Available online at: https://www.nsf.gov/pubs/2007/nsf07547/nsf07547.htm (accessed January 22, 2020).

National Science Foundation (2019). *Dear Colleague Letter: 2019 Social, Behavioral, and Economic (SBE) Repositioning*. Available online at: https://www.nsf.gov/pubs/2019/nsf19089/nsf19089.jsp (accessed April 9, 2020).

Oancea, A. (2013). Buzzwords and Values: The prominence of "impact" in UK research policy and governance. *Res. Trends* 33, 6–9. Available online at: https://www.researchtrends.com/issue-33-june-2013/buzzwords-and-values (accessed July 2, 2020).

Office of Science and Technology Policy (2006). *The Science of Science Policy: A Federal Research Roadmap*. Washington, DC: National Science and Technology Council.

Onken, J. (2016). "Improving the research portfolio data infrastructure at NIH," in *Presentation, Modeling Science, Technology, and Innovation* (Washington, DC). Available online at: https://modsti.cns.iu.edu/wp-content/uploads/2016/05/Onken_DAI.pdf (accessed July 8, 2020).

Pavlopoulos, G. A., Paez-Espino, D., Kyrpides, N. C., and Iliopoulos, I. (2017). Empirical comparison of visualization tools for larger-scale network analysis. *Adv. Bioinform.* 2017:1278932. doi: 10.1155/2017/1278932

Questex (2020). *FiercePharma: The Top 20 Drugs by 2018 U.S. Sales*. Available online at: https://www.fiercepharma.com/special-report/top-20-drugs-by-2018-u-s-sales (accessed February 10, 2020).

Raftery, J., Hanney, S., Greenhalgh, T., Glover, M., and Blatch-Jones, A. (2016). Models and applications for measuring the impact of health research: update of a systematic review for the health technology assessment programme. *Health Technol. Assess*. 20, 1–254. doi: 10.3310/hta20760

RAND Europe. (2008). *Medical Research: What's it Worth? Estimating the Economic Benefits From Medical Research in the UK*. London: UK Evaluation Forum.

Research Excellence Framework (2015). *REF Manager's Report*. Available online at: https://www.ref.ac.uk/2014/media/ref/content/pub/REF_managers_report.pdf (accessed July 2, 2020).

Rigby, J. (2011). Systematic grant and funding body acknowledgment data for publications: new dimensions and new controversies for research policy and evaluation. *Res. Eval*. 20, 365–375. doi: 10.3152/095820211X13164389670392

Ruegg, R., and Jordan, G. (2007). *Overview of Evaluation Methods for R&D Programs: A Directory of Evaluation Methods Relevant to Technology Development Programs*. Report prepared for U.S. Department of Energy Office of Energy Efficiency and Renewable Energy. Available online at: https://www1.eere.energy.gov/analysis/pdfs/evaluation_methods_r_and_d.pdf (accessed January 27, 2020).

Sherwin, C. W., and Isenson, R. S. (1967). Project hindsight: defense department study of the utility of research. *Science* 156, 1571–1577. doi: 10.1126/science.156.3782.1571

Skelton, V. (2019). *International consortium The Research on Research Institute launched. Information Today Europe*. Available online at: https://www.infotoday.eu/Articles/News/Featured-News/International-consortium-The-Research-on-Research-Institute-launched-134441.aspx (accessed October 2, 2019).

Smith, C. I. E., and Zain, R. (2019). Therapeutic oligonucleotides: state of the art. *Annu. Rev. Pharmacol. Toxicol*. 59, 605–630. doi: 10.1146/annurev-pharmtox-010818-021050

Smith, R. (1987). Comroe and dripps revisited. *Br. Med. J*. 295, 1404–1407. doi: 10.1136/bmj.295.6610.1404

U.S. Food and Drug Administration (2020). *Approved Drug Products with Therapeutic Equivalence Evaluations (Orange Book)*. Available online at: https://www.fda.gov/drugs/drug-approvals-and-databases/approved-drug-products-therapeutic-equivalence-evaluations-orange-book (accessed March 15, 2019).

UK Research and Innovation (2020). *The Gateway to Research: UKRI Portal Onto Publically Funded Research*. Available online at: https://gtr.ukri.org/ (accessed July 9, 2020).

van Raan, A. F. J. (2017). Patent citations analysis and its value in research evaluation: a review and a new approach to map technology-relevant research. *J. Data Inf. Sci*. 2, 13–50. doi: 10.1515/jdis-2017-0002

Waldman, L., and Lariviere, V. (2020). Special issue on bibliographic data sources. *Quant. Sci. Stud*. 1, 360–362. doi: 10.1162/qss_e_00026

Williams, R. S., Lotia, S., Holloway, A. K., and Pico, A. R. (2015). From scientific discovery to cures: bright stars within a galaxy. *Cell* 163, 21–23. doi: 10.1016/j.cell.2015.09.007

Wooding, S., Hanney, S., Pollitt, A., Buxton, M., and Grant, J. (2011). *Project Retrosight: Understanding the Returns From Cardiovascular and Stroke Research: The Policy Report*. Santa Monica, CA: RAND Corporation. Available online at: https://www.rand.org/pubs/monographs/MG1079.html (accessed February 6, 2020).

Zuckerman, B. L., Hautala, J. A., and Nek, R. (2015). *Technology Development at the National Institutes of Health (NIH): Summary Report (IDA Document D-5712)*. Washington DC: IDA Science & Technology Policy Institute.

Conflict of Interest: The authors are employed by, or receive contract support from, the National Institute of General Medical Sciences, which funded all aspects of this study.

frontiers in Research Metrics and Analytics

ORIGINAL RESEARCH
published: 16 October 2020
doi: 10.3389/frma.2020.569268

The Scientific Impact Derived From the Disciplinary Profiles

Jüri Allik[1]*, *Kalmer Lauk*[2] *and Anu Realo*[1,3]

[1] *Institute of Psychology, University of Tartu, Tartu, Estonia,* [2] *Grant Office, University of Tartu, Tartu, Estonia,* [3] *Department of Psychology, University of Warwick, Coventry, United Kingdom*

The disciplinary profiles of the mean citation rates across 22 research areas were analyzed for 107 countries/territories that published at least 3,000 papers that exceeded the entrance thresholds for the *Essential Science Indicators* (*ESI*; Clarivate Analytics) during the period from January 1, 2009 to December 31, 2019. The matrix of pairwise differences between any two profiles was analyzed with a non-metric multidimensional scaling (MDS) algorithm, which recovered a two-dimensional geometric space describing these differences. These two dimensions, Dim1 and Dim2, described 5,671 pairwise differences between countries' disciplinary profiles with a sufficient accuracy (stress = 0.098). A significant correlation ($r = 0.81$, $N = 107$, $p < 0.0001$) was found between Dim1 and the *Indicator of a Nation's Scientific Impact* (*INSI*), which was computed as a composite of the average and the top citation rates. The scientific impact ranking of countries derived from the pairwise differences between disciplinary profiles seems to be more accurate and realistic compared with more traditional citation indices.

Keywords: disciplinary profiles, scientific impact, *Essential Science Indicators*, multidimensional scaling, bibliometrics

Edited by:
Juan Ignacio Gorraiz,
Universität Wien, Austria

Reviewed by:
Hamid R. Jamali,
Charles Sturt University, Australia
Zaida Chinchilla-Rodríguez,
Consejo Superior de Investigaciones Científicas (CSIC), Spain

Correspondence:
Jüri Allik
juri.allik@ut.ee

Specialty section:
This article was submitted to Research Assessment, a section of the journal Frontiers in Research Metrics and Analytics

Received: *03 June 2020*
Accepted: *17 August 2020*
Published: *16 October 2020*

Citation:
Allik J, Lauk K and Realo A (2020) The Scientific Impact Derived From the Disciplinary Profiles.
Front. Res. Metr. Anal. 5:569268.
doi: 10.3389/frma.2020.569268

THE SCIENTIFIC IMPACT DERIVED FROM THE DISCIPLINARY PROFILES

Although not perfect, the number of times a scientific paper has been cited since its publication is an objective and easy-to-determine indicator of its scientific impact, which was forecasted long before counting citations became practically feasible (Garfield, 1955). After an expected link between scientific and economic wealth was established—countries whose scientists tend to publish highly cited science papers had also higher level GDP per capita—the mean citation rate acquired a status of the most reliable measure of the scientific quality of nations (May, 1997; Rousseau and Rousseau, 1998; King, 2004; Harzing and Giroud, 2014; Prathap, 2017). However, it was noticed that some countries, for example, Sweden and Finland, seem to have lower mean citation rates than some other countries with a comparable level of scientific development such as Switzerland and the Netherlands (Karlsson and Persson, 2012; Öquist and Benner, 2015). It was also observed that Scandinavian papers published with international co-authorship produced a higher citation rate than purely domestic papers (Glänzel, 2000). It was also noticed that there was a gap between national mean citation rates and the proportion of highly cited papers that countries' scientists were publishing, which could be considered as an index of complaisance showing satisfaction with a relatively modest scientific ambitions (Allik, 2013; Lauk and Allik, 2018).

These and other similar problems caused a shift in the bibliometric research from impact scores based on average values of citations toward the use of indicators that reflect the top of the citation distribution, such as the number of papers reaching the highest rank of citations (van Leeuwen et al., 2003). In accordance with this general trend, a composite index—the *High Quality Science Index* (HQSI; Allik, 2013; Allik et al., 2020a)—was proposed characterizing nations by combining the mean citation rate per paper with the percentage of the papers that have reached the top 1% level of citations in a given research area and an age cohort of published papers. Although the average values of citations and the top of the citation distribution are highly correlated, typically $r = 0.80$ or higher (Allik, 2013), combining these two indicators into a composite index allowed to compensate some minor discrepancies between the two indicators.

Despite these improvements, the rankings of countries based on their citation frequencies are still often counterintuitive, seemingly at least. For example, very few experts would have expected that Panama will become a leading country whose scientists are publishing papers with the highest citation rate in the whole world (Monge-Najera and Ho, 2015; Confraria et al., 2017; Erfanmanesh et al., 2017; Allik et al., 2020a). One possible reason for such implausible rankings is that the selected top layer of papers is not representative of the total scientific production of a given nation (Allik et al., 2020a). When the *Essential Science Indicators* (*ESI*; Clarivate Analytics) database was designed, the whole science (except humanities) was decided to divide into 22 research areas with a quite different publication and citation rates. However, counting minimally required number of citations to enter the *ESI* in one of the research fields created a situation where it may be more advantageous to avoid entering *ESI* in relatively weak research areas that could decrease the country's average citation rate. As was shown by Allik and colleagues (2020a), leaving weaker publications out of counting may artificially increase the mean citation rate of that nation (Allik et al., 2020a). To deal with this problem, a new indicator—the *Indicator of a Nation's Scientific Impact* (*INSI*)—was proposed, wherein, in addition to the average and the top citation rates, the number of research areas in which each country/territory had succeeded to enter the *ESI* was also taken into account (Allik et al., 2020b). This modification made the scientific impact ranking of countries/territories more plausible, unfortunately not entirely. For example, the Republic of Georgia, which had the 5th highest mean citation rate, was shifted five positions down in the ranking because of the failure to exceed the *ESI* entrance threshold in 11 out of 22 scientific fields. However, Panama—also failing in 11 areas—dropped only two positions in the ranking and remained nevertheless ahead of the Netherlands, Denmark, and the United Kingdom not to mention USA, Canada, and Germany.

It was noticed that characteristics of the disciplinary structure may also be a factor that affects the competitive advantages of national sciences (Yang et al., 2012; Bongioanni et al., 2014, 2015; Cimini et al., 2014; Harzing and Giroud, 2014; Radosevic and Yoruk, 2014; Albarran et al., 2015; Lorca and de Andrés, 2019; Pinto and Teixeira, 2020). For example, it has been argued that this archaic disciplinary structure is one of the reasons why Russia and other former communist countries are still lagging behind Western nations (Kozlowski et al., 1999; Markusova et al., 2009; Adams and King, 2010; Guskov et al., 2016; Jurajda et al., 2017; Tregubova et al., 2017; Shashnov and Kotsemir, 2018). In a comprehensive study of how disciplinary structure is related to the competitive advantage in science of different nations, Harzing and Giroud (2014) showed that countries that demonstrated the fastest increase in their scientific productivity during the periods 1994–2004 and 2002–2012 remained relatively stable in their fairly well-balanced disciplinary structures. They also identified different groups of countries with distinct patterns of specialization. For example, one group of countries with a highly developed knowledge infrastructure had an emphasis on social sciences. Another group of countries had a rather balanced research profile with some slight advantage in physical sciences. Yet another group of countries mainly comprised Asian countries with a competitive advantage in engineering sciences (Harzing and Giroud, 2014). Although this study shed light on slightly different routes toward scientific excellence, it is still unclear whether there are truly separate routes or only one general highway, which guarantees advancement in the world ranking.

One of the problems with existing research on examining the scientific disciplinary profiles is that previous studies typically involved a relatively small number of nations. For example, the study by Harzing and Giroud (2014) analyzed disciplinary profiles of 34 countries across 21 disciplines while Almeida et al. (2009) examined disciplinary profiles of 26 European countries. Another study analyzed 27 European countries across 27 disciplines over the period from 1996 to 2011 (Bongioanni et al., 2014). Thelwall and Levitt (2018) analyzed the relative citation impact for 2.6 million articles from 26 fields in the 25 countries published from 1996 to 2015. Pinto and Teixeira (2020) examined disciplinary profiles of 65 countries over a broad period of time (1980–2016). There were several studies analyzing 16 G7 and BRICS countries (Yang et al., 2012; Shashnov and Kotsemir, 2018; Yue et al., 2018). Li (2017) explored disciplinary profiles of 45 countries, which is still a relatively small fraction of nations capable for a substantial scientific contribution. Estimating that there are about 100 nations with sufficiently advanced sciences, a need for more inclusive studies is obvious.

The Aim of the Present Study

To advance the existing research, the aim of the present study is to examine the disciplinary profiles of the mean citation rates for 107 countries or territories whose scientists made substantial contributions to the world's essential science. In accordance with a recommendation to use indicators reflecting the top of the citation distribution (van Leeuwen et al., 2003), we used publications that were selected by the *ESI* based on their top citation rates. For each country/territory that had exceeded the entrance thresholds, their disciplinary profiles were formed based on their mean citation rates across 22 broad disciplines that *ESI* uses to monitor publication and citation performance.

When comparing disciplinary profiles of any two countries, we can judge how similar or dissimilar disciplinary strengths or weaknesses of these two countries are. From pairwise (dis)similarities between any two disciplinary profiles, it is possible to construct a matrix of distances between all countries/territories. By applying a multidimensional scaling (MDS) algorithm to this matrix, we may hope to recover from it a geometric space of low dimensionality, which could represent these data, as they are points in this geometric space (Borg and Groenen, 2005). If axes of this geometric space have a meaningful interpretation, then we may have a novel way for the construction of a new index characterizing the scientific impact of nations, which would not base on the average or the top values of the citation distribution alone.

METHODS

Data were retrieved from the latest available update of the *ESI* (*Clarivate Analytics*, updated on March 12, 2020; https://clarivate.com/products/essential-science-indicators/) that covered an 11-year period from January 1, 2009, until December 31, 2019. This update contained over 16 million *Web of Science* (*WoS*) documents, which were cited over 221 million times with an average frequency of 13.5 times per document.

In order to be included in the *ESI*, journals, papers, institutions, and authors need to exceed the minimum number of citations obtained by ranking journals, researchers, and papers in a respective research field in descending order by citation count and then selecting the top fraction or percentage of papers. For the authors and institutions, the threshold is set for the top 1% and the top 50% is established for countries and journals in an 11-year period. The main purpose of dividing into the fields is to balance publication and citation frequencies in different research areas. The *ESI* entrance thresholds were quite different for the research areas. For example, in the field of clinical medicine, 16,012 citations were needed for a country/territory in order to pass the *ESI* threshold whereas the respective figures in the fields of mathematics and economics & business were 494 and 321.

Among 149 countries/territories that passed the *ESI* threshold at least in one research field were several that published a small number of papers. For example, researchers from the Seychelles, Bermuda, and Vatican published 421, 404, and 257 papers, respectively, which were able to surpass the disciplinary entrance thresholds during the last 11 years. To include countries with a sufficient number of papers, we analyzed only countries that published more than 3,000 papers during the 11-year period. This entrance threshold was slightly lowered compared with the previous studies where it was 4,000 (Allik, 2013; Lauk and Allik, 2018; Allik et al., 2020b) to include a maximally large number of countries/territories making substantial contribution to the world science. Applying this criterion, 107 countries/territories were included in the analyses, which is about 78% of all countries/territories admitted to the *ESI*. The disciplinary profiles for these 107 counties/territories were retrieved from the *ESI*, and the mean citation rates across 22 research areas were reproduced in **Table 1** without any modifications. However, lowering this criterion further to 2,000 would have extended the list by 16 additional countries: Mozambique, Bolivia, Democratic Republic of Congo, Bahrein, Cambodia, Ivory Coast, Jamaica, Madagascar, Yemen, Moldova, Syria, Libya, Mongolia, Trinidad and Tobago, and Montenegro. Because the scientific strength can be measured by the number of disciplines in which a country/territory succeeded to enter *ESI*, we excluded these 16 countries as they succeeded to exceed the entrance thresholds typically only in three to four research areas and no more than in eight areas, which is <40% of the total number of research areas.

Indicator of a Nation's Scientific Impact (*INSI*)

The penultimate column ["*INSI* (rank)"] in **Table 1** presents the country/territory *INSI* score (ranking), which is an average of three components (Allik et al., 2020b). The first component is the country/territory mean citation rate—the number of citations divided by the number of papers (the 4th column from the right "All fields"). The second component is the percentage of papers that had reached the top 1% citation rate in the respective research area and age cohort (the 3rd column from the right "Top 1%"). Finally, the third component is a number of research areas or disciplines in which each country/territory had reached the *ESI* (the number of nonzeros in the first 22 columns). For example, large countries such as USA, Germany, China, and Russia have surpassed the *ESI* entrance thresholds in all 22 research fields. However, 49 (46%) out of 107 countries/territories failed to reach the *ESI* in one research area at least. Before computing the average score, three *INSI* components were normalized so that their mean values were equal to zero with the standard deviation equal to one. Thus, the *INSI* scores in the last column are in the units of the standard deviation showing how much below or above the average score of all 107 countries/territories each participant was scoring.

Table 1 also reproduces the mean citation rates of each country/territory in 22 different research fields. Zeros represent research fields in which country/territory failed to enter the *ESI*. For example, Benin, Bosnia, and Herzegovina, and Uzbekistan had 16 zeros in their disciplinary profiles. Because no entry means no citations, we treated those research areas as if they had zero citation rates.

For the analysis of 107 disciplinary profiles of each country/territory across 22 different research fields, we used the MDS technique, which attempts to transform "distances" or "proximities" among a set of *N* objects into a configuration of *N* points mapped into a geometric space with the smallest possible number of dimensions. A non-metric version of MDS assumes that only the ranks of the distances are known or relevant for producing a map, which reproduces these ranks in the best possible way. We applied the non-metric Guttman-Lingoes MDS algorithm (Borg and Groenen, 2005) as it is implemented in the *Statistica* (*Dell* Inc.) software package. Before applying a MDS algorithm, a matrix of pairwise (dis)similarities between disciplinary profiles of any two countries/territories was computed. The absolute pairwise differences across 22 disciplines were summed together, being used as a measure of (dis)similarity between any pairs of countries/territories. As a result, we created a symmetric matrix with 11,449 elements, each of which showing City Block or Manhattan distance between all possible pairs

TABLE 1 | The mean citation rates in 22 research fields and the average citation rate for 107 countries/territories that published 3,000 or more papers able to enter the *ESI* for the period 2009–2019.

Country/territory	Agricultural sciences	Biology and biochemistry	Chemistry	Clinical medicine	Computer science	Economics and business	Engineering	Environment/ecology	Geosciences	Immunology	Materials science	Mathematics	Microbiology	Molecular biology and genetics	Multidisciplinary	Neuroscience and behavior	Pharmacology and toxicology	Physics	Plant and animal science	Psychiatry/psychology	Social sciences, general	Space science	All fields	Top 1%	*INSI* (rank)	Dim1 (rank)
Panama	14.7	26.0	0.0	99.2	0.0	0.0	0.0	29.6	31.3	0.0	0.0	0.0	14.3	33.5	36.5	22.9	0.0	0.0	15.2	0.0	12.9	0.0	27.3	3.7	2	68
Iceland	12.4	27.9	18.9	31.5	12.1	8.5	11.4	17.5	20.4	24.3	14.2	5.4	16.1	122.8	51.3	27.6	21.9	15.3	12.4	13.2	9.0	51.1	26.4	3.2	1	8
Switzerland	14.5	29.7	24.1	24.9	13.4	12.2	13.1	28.4	24.2	29.9	29.0	6.8	24.7	45.7	44.7	25.8	18.8	24.1	16.5	15.5	12.1	34.6	23.5	2.8	3	1
Georgia	0.0	10.7	14.8	68.3	0.0	0.0	0.0	0.0	10.1	0.0	0.0	2.2	0.0	72.9	0.0	10.5	0.0	32.0	0.0	10.1	0.0	31.1	23.1	4.2	4	98
Netherlands	17.0	25.0	23.8	25.4	10.6	14.2	12.2	24.6	24.8	26.1	25.0	5.9	26.4	44.7	30.1	26.9	18.6	23.9	18.7	18.2	12.1	32.4	22.5	2.5	5	2
Scotland	19.0	28.6	21.0	29.0	11.6	10.2	10.8	23.6	20.1	31.3	19.0	6.3	26.9	47.1	52.3	29.6	24.3	23.4	16.8	16.6	10.2	33.9	22.5	2.7	7	3
Denmark	14.2	26.3	20.5	24.5	9.5	11.6	14.7	23.2	21.5	24.3	23.0	5.7	23.4	44.1	31.4	21.8	19.3	23.2	15.8	16.1	11.5	39.4	21.5	2.5	8	10
Singapore	12.8	23.2	32.7	19.3	14.4	12.3	13.8	19.1	14.2	27.7	35.6	7.0	23.6	43.0	29.3	19.8	19.8	19.4	11.5	13.7	7.7	0.0	21.5	2.7	9	7
Wales	17.9	23.7	19.7	23.9	10.7	10.6	14.1	23.3	20.8	32.2	17.1	5.9	22.7	55.7	19.5	27.9	19.7	21.9	15.0	17.3	10.7	47.0	21.0	2.4	10	4
Estonia	10.8	25.6	15.3	41.8	4.4	5.7	7.5	23.3	12.3	19.9	12.1	4.4	20.8	64.3	20.1	18.6	20.8	24.1	17.0	14.4	6.2	25.0	20.7	2.9	6	26
Belgium	15.3	25.9	19.7	27.5	9.3	11.0	11.9	19.9	20.9	25.1	20.3	5.8	21.2	40.9	42.4	23.5	19.8	20.1	14.7	17.0	9.4	27.2	20.4	2.3	12	11
England	16.0	27.3	21.9	23.4	10.7	11.5	11.3	23.5	22.2	26.1	21.7	6.3	24.6	41.0	35.0	28.2	19.4	20.2	17.0	16.8	10.0	29.4	20.2	2.2	14	5
Ireland	18.0	24.7	21.3	21.5	9.0	9.8	12.8	19.2	18.9	33.0	28.5	5.5	25.0	46.8	40.7	29.4	19.1	19.6	13.1	14.4	8.1	40.0	19.8	2.3	16	12
Sweden	15.5	24.6	19.4	23.1	9.2	11.2	11.8	23.8	21.2	23.4	17.8	5.0	24.2	41.3	33.6	25.4	19.6	18.5	15.8	15.0	10.1	29.1	19.8	2.1	13	6
USA	13.1	26.5	24.7	19.7	11.3	14.2	11.4	19.9	20.5	26.9	28.3	6.5	24.3	37.6	32.9	25.4	18.2	19.9	13.4	16.3	9.7	26.6	19.6	1.8	21	9
N. Ireland	16.3	17.9	21.3	23.9	10.3	9.5	12.5	20.2	34.2	20.2	18.6	9.3	16.5	52.4	23.3	24.1	18.4	18.0	14.1	14.1	7.9	23.9	19.1	2.0	17	13
Austria	14.8	23.1	15.4	22.6	9.5	10.5	9.3	20.8	22.7	26.3	15.6	5.5	23.3	37.5	33.8	23.9	17.6	22.6	13.6	13.9	10.1	28.4	19.0	2.2	15	17
Zambia	0.0	0.0	0.0	36.4	0.0	12.9	0.0	16.2	0.0	19.0	0.0	0.0	17.8	0.0	32.6	0.0	0.0	0.0	8.2	9.7	8.9	0.0	18.8	2.6	40	93
Finland	15.8	23.4	16.5	23.3	10.3	11.2	10.9	20.8	20.6	22.7	15.1	6.3	20.5	46.9	21.5	25.1	18.3	20.5	14.2	14.5	8.8	29.4	18.7	1.9	23	18
Canada	14.4	23.0	19.8	23.6	11.2	12.2	11.4	19.6	17.7	23.0	18.4	5.7	21.7	33.5	25.3	22.8	17.3	19.5	13.7	15.8	10.2	34.5	18.6	1.9	25	15
Germany	12.6	23.6	19.9	19.3	9.8	9.8	9.5	20.1	19.5	25.9	19.7	5.6	21.3	34.3	30.6	22.9	17.3	18.7	15.4	13.7	8.7	29.0	18.6	1.7	27	20
Peru	9.3	13.7	9.3	36.6	0.0	5.3	7.0	21.0	20.7	20.6	9.0	3.6	17.8	39.5	21.2	14.3	10.6	23.3	9.5	13.9	7.6	13.0	18.5	2.8	11	35
Norway	14.3	20.8	14.9	24.3	10.5	10.9	11.4	21.2	20.7	22.1	13.6	6.8	17.9	42.6	35.8	24.4	16.2	19.5	14.2	13.6	10.3	38.0	18.4	2.1	18	19
France	15.1	22.6	18.2	22.0	8.8	9.0	10.2	20.7	19.5	23.9	16.9	5.8	21.1	34.5	28.4	22.3	16.9	17.0	16.3	11.7	8.4	28.8	18.1	1.7	31	22
Australia	13.7	23.6	20.7	20.4	11.4	9.1	13.2	21.3	19.5	25.5	22.1	6.2	20.8	34.6	25.2	21.5	17.5	19.2	14.7	13.9	8.7	28.0	17.7	2.0	24	16
Hong Kong	14.6	20.0	25.0	20.8	12.3	13.9	14.9	18.2	21.3	19.4	26.7	7.8	24.3	29.4	30.9	17.9	16.6	17.7	15.3	13.1	8.4	16.8	17.7	2.2	19	14
Israel	15.0	23.9	20.6	19.0	8.9	9.9	9.4	16.6	17.1	26.3	22.8	5.1	20.4	37.3	32.5	21.6	19.1	19.0	15.3	12.9	7.0	33.5	17.7	1.8	35	38

(Continued)

TABLE 1 | Continued

Country/territory	Agricultural sciences	Biology and biochemistry	Chemistry	Clinical medicine	Computer science	Economics and business	Engineering	Environment/ecology	Geosciences	Immunology	Materials science	Mathematics	Microbiology	Molecular biology and genetics	Multidisciplinary	Neuroscience and behavior	Pharmacology and toxicology	Physics	Plant and animal science	Psychiatry/psychology	Social sciences, general	Space science	All fields	Top 1%	*INSI* (rank)	Dim1 (rank)
Kenya	12.4	19.5	0.0	32.3	0.0	14.4	9.9	19.5	18.2	18.2	0.0	0.0	17.3	33.3	21.4	22.3	10.4	0.0	10.3	13.8	10.8	0.0	17.7	2.2	30	21
Malawi	12.3	71.7	0.0	22.8	0.0	13.9	0.0	0.0	0.0	17.3	0.0	0.0	29.1	48.4	21.4	0.0	0.0	0.0	0.0	10.6	9.8	0.0	17.5	2.1	51	92
Luxembourg	22.8	22.9	16.9	33.6	9.1	6.6	13.7	18.3	19.3	25.1	15.4	4.2	17.2	36.4	11.0	21.7	18.9	14.7	14.3	11.7	8.8	0.0	17.4	2.3	20	23
Uganda	8.2	17.7	0.0	29.6	0.0	8.1	0.0	14.9	13.2	16.5	0.0	0.0	19.7	37.1	10.4	11.0	13.0	0.0	9.1	14.5	9.4	0.0	17.2	1.6	49	70
Italy	13.2	18.2	17.3	20.8	8.6	8.7	11.2	16.2	16.3	23.0	17.3	6.1	16.5	31.0	34.8	21.2	15.2	16.9	11.9	13.7	9.1	27.0	17.1	1.6	28	24
New Zealand	13.1	22.9	15.6	23.3	9.5	9.7	10.8	24.1	20.3	21.3	15.7	4.6	20.5	36.2	14.7	18.8	16.7	21.8	13.1	14.9	8.3	21.3	17.1	1.9	34	25
Sri Lanka	10.7	12.0	12.0	39.6	14.2	5.1	8.0	16.2	9.9	14.0	7.9	0.0	10.3	0.0	7.8	15.0	10.0	24.2	9.0	10.6	7.5	0.0	17.1	2.6	22	55
Costa Rica	11.5	11.8	0.0	32.8	0.0	8.1	0.0	21.9	19.5	22.6	0.0	0.0	11.7	82.0	21.5	21.4	14.3	0.0	9.2	13.6	10.9	0.0	16.6	1.8	45	48
Philippines	11.6	9.7	12.2	42.7	7.4	5.9	11.8	15.3	13.1	17.0	8.0	0.0	13.9	49.6	19.4	13.8	11.6	6.8	11.8	11.7	7.2	0.0	16.5	2.5	26	43
Spain	14.1	20.3	19.0	19.9	8.4	8.9	11.1	17.3	16.2	19.6	16.9	5.3	18.2	34.3	25.7	19.9	14.8	19.5	13.4	10.1	6.9	27.9	16.4	1.5	36	27
Armenia	0.0	0.0	7.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	28.8	0.0	0.0	8.7	23.1	0.0	0.0	0.0	13.2	16.3	2.7	33	28
Benin	8.8	0.0	0.0	55.5	0.0	0.0	0.0	0.0	0.0	14.8	0.0	0.0	17.5	0.0	0.0	0.0	15.9	0.0	8.5	0.0	0.0	0.0	16.3	2.2	77	100
Cyprus	12.1	12.6	17.7	23.9	10.0	8.7	12.3	20.0	21.4	0.0	14.1	5.6	0.0	20.3	0.0	14.0	9.9	30.3	10.2	10.1	6.4	18.0	16.3	2.4	29	31
Greece	13.2	18.7	17.4	19.7	9.3	8.2	12.3	14.3	15.5	25.8	14.5	5.6	17.0	37.4	28.4	16.6	15.1	20.5	11.0	11.6	10.3	21.8	16.3	1.7	54	105
Tanzania	9.6	18.2	0.0	25.1	0.0	8.4	0.0	13.8	13.0	19.1	0.0	0.0	20.9	32.9	19.9	13.0	11.8	0.0	8.5	9.4	10.4	0.0	16.0	1.6	52	67
Portugal	14.2	18.4	16.6	16.4	7.5	8.7	12.1	16.4	16.2	22.7	16.5	4.8	17.3	28.5	20.8	22.5	16.6	21.0	12.1	9.2	7.0	30.8	15.4	1.5	39	29
Hungary	8.0	16.1	12.3	20.5	6.3	6.0	6.8	14.3	11.2	22.4	10.2	3.4	13.4	32.1	21.9	17.5	13.3	22.4	10.1	14.8	7.1	25.4	14.8	1.7	50	41
Uruguay	12.3	17.2	13.8	36.9	4.5	3.6	9.3	18.4	13.5	14.9	0.0	3.2	12.1	17.0	80.1	17.4	11.8	9.4	9.0	20.5	4.5	0.0	14.8	1.3	38	33
Botswana	0.0	0.0	0.0	53.3	0.0	0.0	0.0	10.3	8.6	15.6	0.0	0.0	0.0	0.0	65.2	0.0	0.0	0.0	0.0	0.0	5.5	0.0	14.3	1.4	99	106
Qatar	0.0	11.8	14.5	19.7	11.4	4.9	10.6	9.4	0.0	17.4	13.9	5.0	0.0	30.4	41.9	11.8	7.6	17.4	0.0	5.6	4.8	15.5	14.0	2.3	37	63
Zimbabwe	15.1	0.0	0.0	32.5	0.0	0.0	0.0	12.0	9.3	16.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8.7	7.9	8.1	0.0	13.8	1.8	91	94
Japan	7.8	15.1	15.8	12.9	5.8	5.3	7.3	13.3	15.8	21.8	15.5	3.8	14.3	28.6	22.2	15.7	10.9	12.9	10.6	9.6	7.8	22.6	13.6	0.9	102	97
Sudan	6.9	0.0	7.6	33.2	0.0	0.0	11.8	0.0	17.5	15.4	0.0	0.0	9.8	0.0	0.0	0.0	8.3	0.0	6.2	0.0	0.0	0.0	13.6	1.0	62	37
Ghana	7.0	10.0	8.6	29.9	0.0	5.1	7.3	12.4	10.3	13.8	0.0	0.0	15.5	28.1	8.6	11.1	10.2	0.0	5.9	7.1	7.6	0.0	13.5	1.6	61	83
Slovenia	11.4	21.8	12.2	14.0	7.7	6.2	7.8	15.2	11.1	21.7	11.0	4.6	19.3	32.5	21.5	19.6	13.1	25.5	10.1	12.2	4.1	53.6	13.5	1.4	44	30
Saudi Arabia	9.6	14.1	17.7	12.5	9.4	5.5	10.6	14.9	11.2	16.3	18.3	6.8	10.9	34.2	15.4	12.2	8.4	13.2	11.3	10.6	6.7	15.1	13.4	2.3	32	34
South Africa	8.4	16.2	12.0	23.0	8.5	4.4	9.3	15.1	13.7	19.3	10.6	4.4	16.6	25.7	13.8	17.2	11.2	16.4	10.1	9.1	6.7	31.1	13.4	1.6	42	36
Lebanon	11.3	10.8	10.0	19.8	5.8	10.3	8.8	10.1	7.4	16.3	11.3	2.5	10.7	20.3	0.0	11.9	12.3	8.8	8.0	23.8	9.2	45.1	13.3	1.8	41	49

(Continued)

TABLE 1 | Continued

Country/territory	Agricultural sciences	Biology and biochemistry	Chemistry	Clinical medicine	Computer science	Economics and business	Engineering	Environment/ecology	Geosciences	Immunology	Materials science	Mathematics	Microbiology	Molecular biology and genetics	Multidisciplinary	Neuroscience and behavior	Pharmacology and toxicology	Physics	Plant and animal science	Psychiatry/psychology	Social sciences, general	Space science	All fields	Top 1%	*INSI* (rank)	Dim1 (rank)
Czech Republic	8.4	13.5	13.2	22.7	5.4	3.8	6.7	16.3	10.9	17.0	10.0	4.2	13.3	25.8	14.9	18.0	13.5	16.3	10.5	8.5	4.6	17.0	13.2	1.3	83	85
Nepal	9.5	15.4	0.0	16.4	0.0	8.1	10.1	12.6	20.0	14.5	19.2	0.0	0.0	0.0	0.0	0.0	11.7	0.0	6.6	10.7	7.8	0.0	13.2	1.6	48	44
Chile	9.2	11.6	9.5	16.5	7.4	4.9	9.9	12.9	15.2	11.2	8.8	5.0	11.9	27.3	16.1	17.4	12.2	16.7	8.5	7.0	4.5	27.6	13.1	1.4	47	45
Malta	0.0	23.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	62.1	0.0	16.6	12.5	0.0	0.0	6.1	4.9	24.3	13.0	1.8	94	102
Argentina	10.6	12.6	10.9	22.6	5.7	5.8	8.8	14.7	10.8	18.0	10.7	4.0	10.6	19.6	18.0	18.9	10.9	17.0	8.9	9.4	4.9	15.1	12.8	1.1	56	46
Burkina Faso	7.9	0.0	0.0	0.0	0.0	0.0	17.1	0.0	0.0	17.7	0.0	0.0	15.8	27.4	21.9	0.0	0.0	0.0	8.7	0.0	8.9	0.0	12.5	1.0	104	96
Latvia	10.0	13.2	9.0	28.7	0.0	7.4	4.4	16.0	0.0	19.7	8.8	0.0	0.0	74.4	0.0	0.0	9.7	8.3	8.0	0.0	6.8	0.0	12.5	1.9	72	90
Bulgaria	6.0	6.6	9.3	20.0	4.9	6.8	7.6	12.4	11.6	16.0	8.6	3.7	12.1	37.7	1.4	14.7	8.7	23.0	7.3	25.9	6.9	14.0	12.4	1.5	73	32
Taiwan	13.1	14.5	15.2	13.2	8.1	7.1	9.0	11.2	14.2	13.6	15.7	4.8	12.7	18.6	14.4	13.8	13.5	14.0	10.9	9.6	7.5	28.1	12.4	0.8	46	54
Colombia	7.9	10.9	10.2	20.3	6.8	5.8	6.1	16.4	13.4	16.3	8.9	2.8	12.5	27.4	12.9	14.9	8.4	21.2	7.0	11.1	5.5	14.9	12.3	1.7	43	52
South Korea	8.3	14.7	16.7	11.6	6.1	6.6	7.6	10.5	13.7	15.1	16.7	4.5	10.8	19.1	17.6	13.9	13.0	12.6	9.9	9.4	6.9	18.4	12.2	0.9	71	40
Belarus	0.0	10.0	5.5	0.0	0.0	0.0	3.7	0.0	0.0	0.0	5.3	2.4	0.0	24.9	0.0	0.0	11.0	20.7	0.0	0.0	0.0	0.0	12.0	2.0	92	99
Cameroon	7.3	10.2	7.9	23.8	0.0	4.1	7.7	17.1	7.3	12.7	9.6	3.4	14.7	32.7	15.0	13.6	10.0	5.8	7.6	0.0	7.1	0.0	11.9	1.3	70	77
Oman	9.4	11.7	12.0	23.2	5.1	4.3	9.7	10.3	9.4	0.0	9.1	3.3	15.3	21.7	0.0	12.2	10.3	10.1	11.6	8.9	0.0	0.0	11.9	1.5	74	76
Bangladesh	8.2	9.9	9.2	28.8	5.6	5.1	9.1	13.2	9.9	18.1	7.1	0.0	15.5	22.5	8.8	13.0	10.4	7.2	6.1	8.2	8.9	0.0	11.8	1.4	67	47
Croatia	10.1	12.5	11.0	13.5	5.2	3.1	5.6	11.0	10.1	18.2	7.3	3.7	14.4	59.1	42.3	16.2	11.1	23.4	6.0	8.4	3.6	23.6	11.8	1.3	60	69
Thailand	9.6	12.4	11.6	15.3	6.3	6.4	10.3	10.9	12.0	17.8	10.8	3.9	13.4	19.9	20.2	14.6	11.9	13.8	8.7	9.5	7.0	13.3	11.8	1.0	55	51
China	10.2	12.1	15.5	9.9	8.0	7.3	8.9	11.2	12.0	12.8	15.7	5.0	10.3	14.2	17.1	12.3	10.3	10.2	10.0	8.4	8.3	14.0	11.7	1.2	98	91
Indonesia	7.7	13.2	10.0	27.0	4.6	5.1	9.6	17.0	12.1	15.0	7.2	3.1	15.5	14.5	14.9	12.8	8.1	5.2	8.1	7.2	7.7	0.0	11.7	1.2	66	62
Senegal	9.1	0.0	0.0	0.0	0.0	6.2	0.0	12.0	13.8	18.2	0.0	3.0	14.2	26.7	34.4	0.0	0.0	0.0	10.5	0.0	6.8	0.0	11.7	1.2	59	42
Venezuela	5.5	17.8	13.1	15.2	9.3	1.9	6.5	15.1	16.2	19.3	8.3	3.3	14.8	40.0	44.5	13.8	9.9	8.9	7.1	15.8	6.5	20.3	11.6	1.1	65	39
UAE	8.4	11.4	15.2	18.6	6.5	5.9	9.4	14.7	9.7	19.4	10.5	4.5	13.9	22.8	20.7	16.5	9.8	7.6	8.2	6.1	5.7	13.0	11.3	1.4	53	50
Ecuador	7.0	11.6	6.3	21.6	5.2	4.0	5.7	16.0	13.4	25.4	10.8	0.0	11.0	16.9	0.0	12.4	9.3	13.9	7.3	0.0	4.4	0.0	11.2	1.8	57	53
Malaysia	10.0	13.1	12.5	12.7	8.2	5.4	13.3	12.5	11.0	14.6	10.7	4.7	10.7	14.3	4.6	10.3	10.1	11.5	7.6	8.4	6.0	13.3	11.2	1.3	63	80
Slovakia	7.6	10.2	10.0	19.4	5.3	2.5	5.6	12.2	8.5	17.4	6.4	4.6	10.7	26.1	15.1	16.2	14.1	16.4	8.8	5.0	6.9	10.4	11.1	1.1	69	58
Azerbaijan	0.0	0.0	3.6	0.0	8.2	0.0	5.3	0.0	0.0	0.0	3.5	2.4	0.0	0.0	0.0	0.0	0.0	20.8	0.0	0.0	0.0	0.0	11.0	2.0	100	103
Ethiopia	7.2	7.4	6.7	19.9	0.0	10.6	7.3	8.9	13.0	11.1	0.0	5.3	9.4	23.0	5.8	12.2	10.3	0.0	7.0	9.7	8.3	0.0	11.0	1.3	64	88
Macau	0.0	19.8	11.9	0.0	11.3	7.4	9.4	0.0	0.0	0.0	17.0	3.3	0.0	45.1	22.7	7.0	10.6	0.0	0.0	6.2	4.9	7.2	11.0	2.3	80	79
Cuba	7.1	12.2	8.5	0.0	7.5	0.0	7.0	10.4	12.1	16.3	8.5	0.0	17.3	0.0	0.0	16.4	8.8	16.8	6.7	8.4	7.0	0.0	10.6	0.8	97	82

(Continued)

TABLE 1 | Continued

Country/territory	Agricultural sciences	Biology and biochemistry	Chemistry	Clinical medicine	Computer science	Economics and business	Engineering	Environment/ecology	Geosciences	Immunology	Materials science	Mathematics	Microbiology	Molecular biology and genetics	Multidisciplinary	Neuroscience and behavior	Pharmacology and toxicology	Physics	Plant and animal science	Psychiatry/psychology	Social sciences, general	Space science	All fields	Top 1%	*INSI* (rank)	Dim1 (rank)
Mexico	8.7	12.1	9.6	18.4	6.2	4.8	7.0	10.7	10.5	14.9	8.0	3.6	11.2	21.4	21.7	11.9	10.1	13.2	7.4	9.7	5.7	17.7	10.6	1.0	76	57
Lithuania	6.2	12.9	10.4	13.9	6.1	6.5	4.9	10.4	10.4	25.6	6.5	2.9	14.4	45.1	0.0	11.5	10.6	17.0	6.1	11.2	5.5	18.5	10.4	1.3	68	56
India	6.1	11.7	12.3	11.4	5.8	6.4	8.6	11.3	9.5	12.4	11.0	4.0	10.2	13.7	6.8	11.0	11.2	10.1	7.7	9.1	6.3	15.4	10.3	0.7	79	65
Morocco	9.4	9.9	9.9	12.3	4.3	5.1	6.6	9.1	11.8	21.0	9.5	3.3	10.2	15.4	0.0	7.3	11.2	16.5	7.7	11.7	8.9	21.4	10.3	1.1	78	60
Poland	8.3	11.3	9.8	15.1	6.4	3.5	6.0	8.2	8.3	17.1	8.0	3.8	9.1	19.0	11.2	13.6	10.6	12.8	6.7	8.6	5.7	26.1	10.3	0.9	86	64
Vietnam	8.1	10.4	8.5	31.6	7.4	4.2	7.3	9.8	11.4	19.3	8.1	3.4	18.1	22.2	5.9	13.6	7.7	9.4	7.1	6.3	8.4	16.0	10.0	1.5	58	61
Jordan	7.4	9.6	8.8	18.2	5.6	6.4	8.0	7.2	9.8	14.9	8.4	7.0	0.0	24.6	0.0	12.8	7.0	6.8	6.3	7.6	5.1	0.0	9.9	1.2	85	81
Kuwait	10.4	11.8	11.3	11.5	9.9	4.3	8.9	9.8	8.1	21.2	9.5	3.3	10.0	24.1	0.0	9.7	10.3	0.0	0.0	6.8	5.4	0.0	9.9	1.4	82	78
Brazil	6.3	11.8	10.9	11.8	6.2	4.3	7.8	11.7	11.4	14.6	9.2	4.2	11.1	14.7	10.1	14.2	10.1	13.3	6.3	10.9	5.4	20.5	9.8	0.7	87	59
Serbia	8.1	6.6	8.6	10.8	5.3	2.8	5.9	8.2	9.9	15.8	8.4	6.8	10.2	28.4	0.0	13.7	9.4	21.0	5.7	9.2	3.7	9.6	9.6	1.1	81	73
Egypt	7.3	8.3	10.6	10.5	5.4	6.6	8.4	9.4	6.7	12.7	10.8	5.7	8.7	13.4	33.6	11.3	9.3	12.0	7.1	8.4	6.3	5.7	9.5	0.8	88	66
Iran	8.3	8.8	11.7	8.5	7.5	6.6	9.3	8.7	8.4	8.7	11.7	4.2	6.0	10.0	4.6	9.8	9.2	9.3	5.5	6.9	5.3	12.1	9.2	0.8	89	75
Pakistan	7.5	10.0	8.9	12.1	6.0	4.6	8.6	10.5	8.3	16.7	8.4	5.1	11.1	23.7	4.9	12.6	5.9	12.5	5.4	7.4	6.5	10.2	9.0	1.2	75	71
Nigeria	6.1	7.9	9.3	14.3	4.1	4.7	7.7	6.9	5.1	11.4	8.1	3.0	9.4	33.7	8.4	11.0	6.5	6.5	5.4	13.6	4.6	6.0	8.8	1.0	84	84
Ukraine	9.5	11.5	7.4	46.6	4.2	0.7	4.1	10.6	8.0	13.0	4.7	2.8	7.3	12.3	12.6	6.3	10.9	9.9	8.6	13.0	6.6	11.2	8.6	0.9	90	86
Tunisia	11.2	10.0	8.3	11.6	4.2	7.0	7.0	10.7	7.8	13.3	8.5	3.5	9.2	12.3	13.1	8.5	10.8	6.9	8.5	8.2	7.6	0.0	8.5	0.5	96	74
Turkey	8.3	8.8	11.0	6.8	8.8	6.7	10.0	6.8	10.6	15.7	9.4	4.6	6.8	19.6	4.9	8.3	8.6	13.0	5.1	5.7	4.0	15.0	8.3	0.6	93	72
Algeria	8.1	8.9	7.5	28.4	5.0	0.0	8.5	6.3	7.5	12.3	8.7	3.8	10.2	0.0	0.0	15.8	6.9	6.1	7.5	0.0	0.0	32.6	8.1	0.8	101	89
Iraq	5.7	6.8	8.0	22.6	4.1	0.0	7.0	0.0	4.8	0.0	6.5	3.8	0.0	0.0	0.0	8.4	5.8	4.7	0.0	28.0	0.0	0.0	7.9	1.1	103	95
Bosnia and Herzeg	0.0	0.0	0.0	0.0	6.2	0.0	4.2	0.0	0.0	0.0	0.0	2.6	0.0	0.0	0.0	0.0	0.0	16.7	0.0	5.7	0.0	0.0	7.6	1.2	106	104
Russia	6.7	9.7	6.2	16.0	2.9	4.5	4.0	7.9	6.8	16.1	6.0	2.8	10.1	13.9	14.4	7.3	7.8	8.6	6.5	4.4	3.5	13.6	7.6	0.6	95	87
Uzbekistan	0.0	0.0	3.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.3	3.1	0.0	0.0	0.0	0.0	0.0	6.3	0.0	0.0	0.0	18.0	6.7	0.5	107	107
Kazakhstan	0.0	7.7	4.4	0.0	5.3	0.0	3.3	0.0	7.0	0.0	4.9	2.4	0.0	0.0	33.9	0.0	0.0	8.1	0.0	0.0	0.0	14.5	6.5	0.7	105	101

*Zeros indicate research fields in which a country/territory did not exceed required number of citations needed to enter the ESI. The penultimate column shows the Indicator of a Nation's Scientific Impact (INSI) ranking, which, in equal parts, takes the mean citation rate, percentage of papers reaching top 1% citation rate, and the number of research fields that have passed the entrance threshold. The last column shows ranking on the x-axis (Dim1) of **Figure 1**. Countries are ranked according to the mean citation rate (the 4th column from the right "All fields").*

of countries/territories including the main diagonal of zeros representing distance from oneself. In order to compress the large range of differences in these (dis)similarities, we normalized (the mean value became zero with the standard deviation one) sums of absolute differences across 107 countries/territories.

RESULTS

Table 1 presents the mean citation scores across 22 research fields for each country/territory. Entries in the table are ranked according to the mean citation rate (the 4th column from the right "All fields"). Panama (27.3), Iceland (26.4), and Switzerland (23.5) had the highest mean citation rate (note that fields with zeros were not used for the calculation of the mean citation rate). Among 107 countries/territories, publications authored by Russian (7.6), Uzbekistan (6.7), and Kazakhstan (6.5) scientists had the smallest impact on the world science in terms of cited work.

To obtain a better impression about the disciplinary profiles, **Figure 1** displays the mean citation rates across 22 research areas for Switzerland, USA, and China. These three nations were chosen for illustrative purposes only: China and USA were two the most prolific sciences in the world publishing over 2 million and 4 million papers, respectively, during the observed 11-year period; Switzerland was one of the most efficient sciences by the mean citation rate. Please notice that these three nations entered the *ESI* in all 22 disciplines. The *ESI* average citation rate in each research area is also shown as a black broken line providing a baseline with which each nation can be compared. Switzerland, as a long-time efficiency front-runner, has a higher citation rate than USA in almost all research areas. Although the impact of Chinese science is growing (Leydesdorff and Wagner, 2009; Leydesdorff et al., 2014), its mean citation rate is still below the *ESI* average in almost every 22 research fields.

The penultimate column in **Table 1** presents the *INSI* ranking for each country/territory. According to this ranking, the highest-quality science is produced in Iceland, Panama, Switzerland, Republic of Georgia, Scotland, Estonia, Netherlands, Singapore, Denmark, and Wales. As these are all relatively small countries, this confirms previous findings that small countries seem to have an advantage in publishing high-impact scientific papers (Allik et al., 2020a,b). According to the *INSI*, the smallest impact among these 107 countries/territories had publications authored by researchers from Iraq, Burkina Faso, Kazakhstan, Bosnia and Herzegovina, and Uzbekistan. It is also interesting to notice that China, in spite of the increasing research volume, occupied a position in the middle of the *INSI* ranking (the 59th position) and Russia was very close to the bottom (the 95th position).

Next, we were interested in how the disciplinary profiles of the mean citation rates were related to the overall scientific impact of countries/territories. In a previous study (Bongioanni et al., 2014), a complex index, borrowed from the physics of magnetism, was proposed to estimate overlaps between disciplinary profiles of countries. In this study, we preferred a simpler approach computing the sum of absolute differences across all 22 fields between two disciplinary profiles (see *Methods* section). The findings showed that a two-dimensional solution was optimal (stress function = 0.098), showing that all differences between countries/territories can be placed on a plane with a sufficient accuracy (cf. Mair et al., 2016).

Figure 2 shows a two-dimensional plot derived from the MDS of similarities–differences between the disciplinary profiles of the mean citation rates. The first dimension Dim1 can be identified as the country/territory's overall scientific impact. Rankings of countries/territories on this dimension Dim1 is presented in the last column in **Table 1** ["Dim1 (rank)"]. The correlation between Dim1 and the *INSI* was $r = 0.81$ ($N = 107$, $p < 0.0001$), which is higher than correlations between Dim1 and any of the *INSI* three components: the mean citation rate ($r = 0.64$, $p < 0.00001$), the percentage of the top-cited papers ($r = 0.35$, $p < 0.00001$), and the number of areas represented in the *ESI* ($r = 0.77$, $p < 0.00001$). After excluding two largest outliers—Panama and Georgia—the correlation increases to $r = 0.88$. Thus, this indicated that the scientific impact of nations could be measured using the Dim1 scores with approximately the same accuracy as with the *INSI*. It is important to emphasize that this ranking was obtained by ignoring the absolute mean citation rates, which is the foundation of the *INSI*. When a transformation for the pairwise differences AB, AC, and BC between any triples of the disciplinary profiles A, B, and C were searched to satisfy an approximate equality $AB + BC \approx AC$, information about the absolute elevation of profiles was lost. Because the triangulation rule was sustained with a reasonable accuracy, it indicated that all differences between profiles can be arranged on a linear scale.

To illustrate how this derived ranking of the scientific impact [**Table 1**, the last column "Dim1 (rank)"] has certain advantages before the previous ones, we need to observe changes in the ranking positions of countries, whose high positions may not be entirely justified. According to the mean citations rate, by which countries/territories were listed in **Table 1**, Panama (27.3) was number one in the world, Georgia (23.1) was on the 4th position and Peru (18.5) occupied the 22nd position. Because the *INSI* penalizes for failures to reach the *ESI* in any of the research areas, countries/territories not being successful in all 22 research areas were expecting to lose positions in the ranking. Because both Georgia and Panama did not enter the *ESI* in 11 research areas, they were shifted down in the *INSI* ranking to the 2nd and the 5th positions, respectively. At the same time, Peru did not reach the *ESI* only in one field (computer sciences); the position in the *INSI* ranking was elevated up to the 12th position.

Compared with these relatively small changes in the ranking of countries that was based on either the mean citation rate or the *INSI*, the differences in the countries' ranking positions on Dim1 derived from the MDS analysis were more substantial. According to their positions on Dim1 (the last column in **Table 1**), Georgia, Panama, and Peru occupied the 98th, 68th, and 35th positions, respectively. Thus, in comparison with the mean citation rate, Georgia, Panama, and Peru dropped 94, 67, and 13 positions. Their disciplinary profiles were more similar to the disciplinary profiles of nations in vicinity to these positions. Two countries with profiles the most similar to Panama in this new ranking were

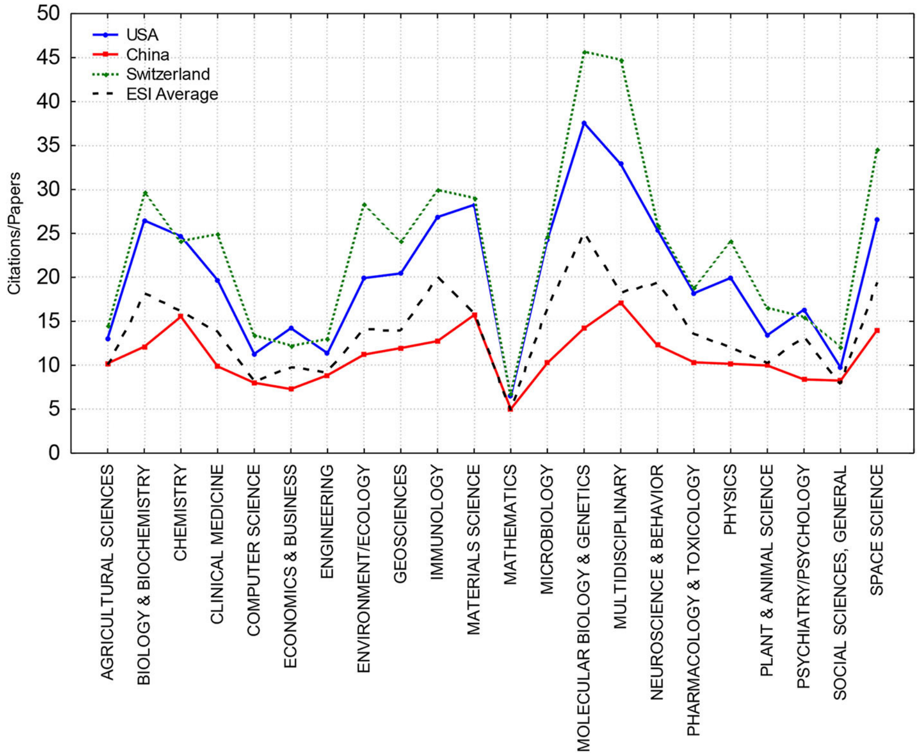

FIGURE 1 | The mean citation rates across 22 research fields for three countries, USA (blue), China (red), and Switzerland (green), compared with the *ESI* average (black broken line).

Tanzania and Bangladesh. Georgia was squeezed between Sudan and Belarus, not Belgium and Ireland as previously in the *INSI* ranking. These changes in the ranking positions explained, as was already mentioned, a relatively modest correlation between Dim1 and the mean citation rate ($r = 0.64$).

The second dimension Dim2 was more difficult to interpret because a clear pattern did not emerge. Because it has the largest positive correlations with the mean citation rate in Clinical Medicine ($r = 0.49$, $p < 0.00001$) and Social Sciences ($r = 0.50$, $p < 0.00001$) to the contrast negative correlations with the mean citation rates in Physics ($r = -0.49$, $p < 0.00001$) and Mathematics ($r = -0.49$, $p < 0.00001$), it would be fair to say that this dimension represents human-centered opposite to the physics–math-centered sciences.

Two distinct clusters can be identified on the plot. These two clusters, we need to warn, were identified based on an impression with a heuristic purpose only, not in the result of any rigorous procedure. The first cluster (surrounded by the red circle) represents the cream of the crop in the world of science. This cluster of 29 countries includes mainly European countries such as the Netherlands, Scotland, and Switzerland but also other countries such as USA, Singapore, Hong Kong, Canada, Australia, Israel, and New Zealand. Although it was noticed that the scientific wealth of Hong Kong and Singapore is declining (Horta, 2018), they firmly belong to this group of leaders in the world science. A common feature of these 29 countries/territories is that they all succeeded to pass the *ESI* entrance thresholds in all 22 research areas.

Another group (green circle) unites not only many of the world's largest countries—China, Russia, Brazil, and India—but also smaller countries like Slovenia, Ecuador, and Hungary. If large countries in this cluster were successful in all 22 disciplinary areas, then smaller countries may have difficulties to collect enough citations to exceed the *ESI* entrance thresholds in some research areas. Outside of these two groups (or circles) are mainly African countries (upper part) or post-communist countries (lower part), which scatter along Dim2.

A similarity between these two clusters and two clusters that were identified previously (Bongioanni et al., 2015, Figure 2) can be noticed. Bongioanni et al. (2015) identified a cluster that included countries with a prominent biomedical disciplinary profile such as the US and the Netherlands (Bongioanni et al., 2015). Another cluster embraced a group of countries with a conspicuous physical-sciences profile, like China and Russia. In addition, many Central, Southern, and Eastern European countries belonged to this second group, as well as India, Indonesia, and Mexico. However, there are notable differences between the findings of Bongioanni et al.'s (2015) and this study. According to Bongioanni et al. (2015), Turkey is in the same group with the UK and the Netherlands; in the current study, Turkey's nearest neighbors are Serbia and Iran in the second group. In addition, Estonia and Portugal were differently

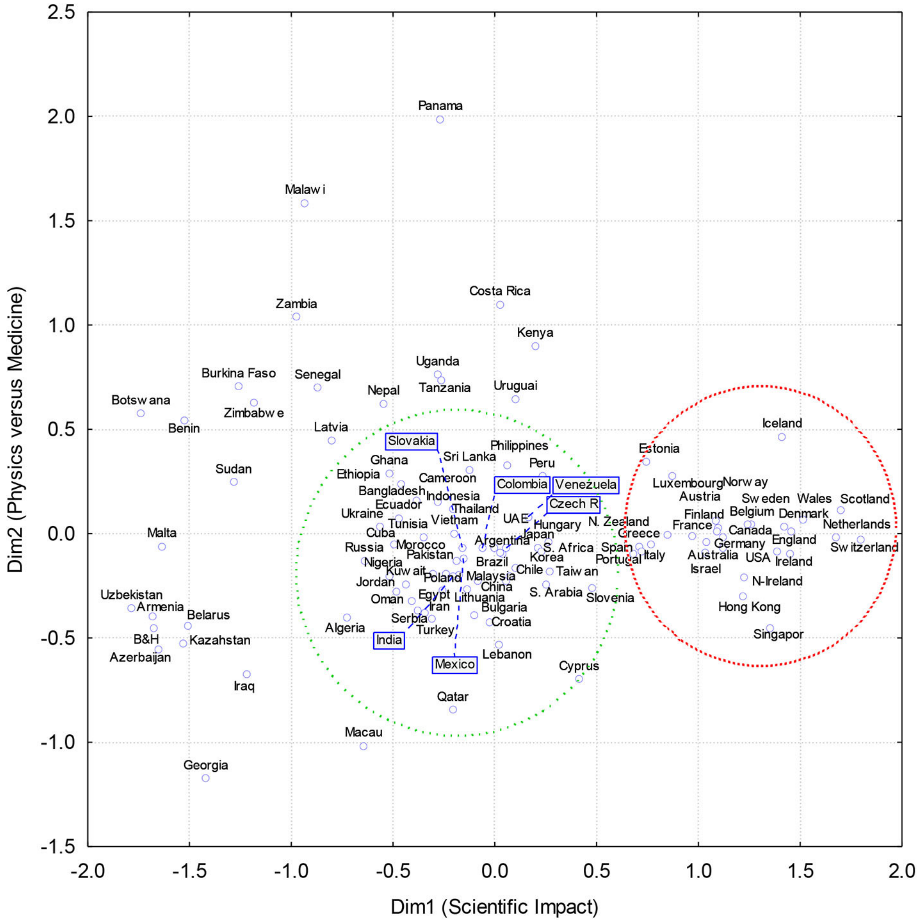

FIGURE 2 | Two-dimensional plot of a non-metric Guttman-Lingoes multidimensional scaling analysis of country's citation profile similarities (Manhattan or City Block metrics of the normalized citation rates).

classified. According to Bongioanni and colleagues (2005), these two countries are in the less scientifically advanced group of nations, while in our classification, they more likely belong to the leading group science nations (Bongioanni et al., 2015). These discrepancies are probably produced by different measures of (dis)similarity between disciplinary profiles.

DISCUSSION

It has been suggested by experts that new impact indicators should not be introduced unless they have a clear added value relative to the existing indicators (Waltman, 2016). Indeed, the average citation rate or the percentage of papers reaching the top of the citation distributions have proved to be trusted and reliable indicators of the scientific wealth of nations (May, 1997; Rousseau and Rousseau, 1998; van Leeuwen et al., 2003; King, 2004; Halffman and Leydesdorff, 2010; Prathap, 2017). Very serious arguments are needed to introduce yet another indicator. Although warnings are still released not to take citations as the only constituents of the concept of scientific quality (MacRoberts and MacRoberts, 2018; Aksnes et al., 2019), citation indicators have become the most convenient measures of the scientific strength of nations (May, 1997; Rousseau and Rousseau, 1998; King, 2004; Harzing and Giroud, 2014; Prathap, 2017). Nevertheless, some of the country rankings based on the citation statistics did not look credible (Allik, 2013; Allik et al., 2020a,b). One of the possible causes of these counterintuitive rankings, as was mentioned above, appears to be the selectivity of databases, which is the main tool for extracting what is believed to be essential in science (Allik et al., 2020a,b). Although it appears to be true that the top of the citation distribution is a more informative characteristic of the scientific impact than indicators based on average values (van Leeuwen et al., 2003), the selectivity of databases unwillingly eliminates "losers" whose counting would have decreased the mean citation rate. Thus, the scientific impact of nations can be increased not only by the number of highly cited papers but also by neglecting those papers that could jeopardize the mean citation rate (Allik et al., 2020a). To improve citation indicators, a new measure—*INSI*—was proposed, which, in addition to the citation statistics, takes also into account the number of research areas in which a country/territory was successful to enter the *ESI*. This amendment improved rankings in the right direction, unfortunately not radically enough (Allik et al., 2020b). As we said above, the disciplinary profiles appeared to be different for scientifically developed and non-leading countries (Kozlowski et al., 1999; Almeida et al., 2009; Yang et al., 2012; Bongioanni et al., 2014, 2015; Harzing and Giroud, 2014; Carley et al., 2017; Li, 2017; Daraio et al., 2018; Shashnov and Kotsemir, 2018). For example, it was noticed that one of the reasons why post-communist countries are still lagging behind Western counterparts is their archaic disciplinary structure reflecting, among other things, the demands of the former totalitarian regimes (Kozlowski et al., 1999; Markusova et al., 2009; Jurajda et al., 2017). Openness of national science systems was observed to be correlated with the scientific impact—the more internationally engaged a nation is, in terms of coauthorships and researcher mobility, the higher the impact on their scientific work (Wagner et al., 2018). It was noticed that geographical proximity, which is one of the strongest incentives for cooperation, may be a principal factor of the similarity between disciplinary profiles (Almeida et al., 2009). Although such pairs as Finland–Norway, England–Scotland, Netherlands–Belgium, and Denmark–Sweden (Almeida et al., 2009, Figure 4) support this idea, there is an equally large number of neighboring countries (e.g., Panama–Colombia, Peru–Ecuador, Georgia–Armenia, Estonia–Latvia, etc.) that have a distinctly different level of scientific impact. It was also observed that BRICS countries differ from the scientifically leading countries typically belonging to G7 not only by the overall scientific impact but also by differences in the disciplinary structure of their sciences (Bornmann et al., 2015; Li, 2017; Shashnov and Kotsemir, 2018; Yue et al., 2018). For example, it was observed that a competitive advantage of a group of nations including the Netherlands, USA, UK, Canada, and Israel is an emphasis on social and biomedical research (Harzing and Giroud, 2014). The disciplinary citation profiles of G7 and BRICS countries are noticeably different. For instance, most G7 countries performed well in Space Science, which was not the strength of BRICS countries (Shashnov and Kotsemir, 2018; Yue et al., 2018). In spite of these differences, there seems to be a common evolutionary pattern of convergence in the national disciplinary profiles (Bongioanni et al., 2014; Bornmann et al., 2015; Li, 2017).

Typically, the disciplinary profiles were analyzed to discover different clusters into which nations belong. Another approach, adopted in this study, was to see if there is a small number dimensions that can summarize (dis)similarities between the disciplinary profiles (cf. Borg and Groenen, 2005). It is not likely that the similarities and dissimilarities between disciplinary profiles have a distinct pattern, which could be described by a low-dimensional space. Like any other human enterprises, science is a complex institution, which may have differences in prioritizing various research fields. For example, Panama in collaboration with the Smithsonian Institution—one of the world's largest museum, education, and research complexes—invested into the study of the tropical ecosystems by creating a branch of the Smithsonian in Panama, which attracted the best researchers around the world in this area (cf. Rubinoff and Leigh, 1990). Another already mentioned example is Georgia allocating considerable assets into physics in order to develop partnerships with the large international collaborative networks. As a result, Georgia achieved the highest mean citation rate (on average 32 cites per paper) in physics (see **Table 1**, column "Physics"). Inspecting **Table 1**, one can also notice, with a surprise, that Kenya had the highest impact among 107 nations in economics and business: every paper that was published by Kenyan's economists collected 14.4 citations on average (column "Economics and business"). Kenya benefited from the research unit of the United Nations Environment Programme in Nairobi, which is devoted to the study of the economics of ecosystems management and provided services (cf. Ivanova, 2007). Knowing the accomplishments that the deCODE and Kári Stefánsson with his colleagues (Hakonarson et al., 2003) have achieved, it is not surprising that Iceland seized the first position in the impact ranking in the molecular biology and genetics (column "Molecular biology and genetics"). These examples seemed to suggest that nations might have different keys for their success in producing high-quality science.

Nevertheless, all (dis)similarities between disciplinary profiles can be arranged on a single dimension ranking, which corresponded to the scientific impact that was measured by conventional indicators such as the *INSI*. This demonstrated that in spite of differences in the nations' competitive advantages, all that mattered was overall impact across many disciplines as possible, not how this impact was allocated among various research areas. To attain success, it was essential to have an evenly high level of citations relative to the *ESI* average across as many disciplines as possible because low impact or not even exceeding the entrance thresholds in one or several research areas is a key factor that diminishes scientific impact. This may also demonstrate that attempts of the agencies that fund scientific research in prioritizing their disciplinary budgets are not as effective as usually claimed. Results of this study appeared to suggest that the only thing that was really worth prioritizing is the scientific excellence irrespective of which particular discipline it was demonstrated. To our satisfaction, the impact ranking derived from the (dis)similarities between disciplinary profiles was free from anomalies that traditional citation indicators typically possess. These results support an idea about a common route toward scientific excellence in which disciplinary peculiarities are supporting a general advancement (Bongioanni et al., 2014; Li, 2017; Thelwall and Levitt, 2018).

In conclusion, previous attempts to construct indicators of the scientific impact of nations were based on the average or the top-citation statistics. However, the country rankings based on these indicators often look problematic and counterintuitive. Most of these anomalies were produced by failures to exceed the *ESI* entrance thresholds in weaker research areas in which nations failed to collect a sufficient number of citations (Allik et al., 2020a,b). To correct these implausible rankings, we proposed to take also into account the number of research areas in which each country/territory failed to exceed the *ESI* entrance thresholds (Allik et al., 2020b). This was an improvement that, however, did not eliminate problematic rankings entirely. In this study, we proposed a novel approach according to which the scientific impact can be derived from the MDS analysis of (dis)similarities between the disciplinary profiles of the mean citation rate. The scientific impact was derived from a matrix of (dis)similarities between disciplinary profiles as a dimension of a recovered geometric space, which characterized the quality of sciences surprisingly adequately without artificially increasing the impact by withdrawing data in weaker research areas. Because shapes of the disciplinary profiles seemed to be irrelevant, only the cumulative citation rate across all disciplines matters in achieving a position in the science impact ranking.

There are several limitations in this study. The decision to include countries that were able publish 3,000 (instead of the previously used 4,000) or more papers during the 11-year period was a voluntary decision. However, some tests with a different number of countries demonstrated that the final plot of the MDS was invariant to this number and preserved its general configuration. Another potentially problematic decision was to replace unrepresented fields with the zero citation rates. We can only guess what the replacement zeros with the actual citation frequencies, which are expectedly close to nil anyway, would have resulted. Unfortunately, the *ESI* does not provide information about the number of publication and their citation rates that were left behind the entrance thresholds. Although we are among the first who noticed that the problem of spurious country rankings may be created by the *ESI*'s most precious property—focusing exclusively on the top of the citation distribution—we have very little information that the application of MDS to the disciplinary profiles provides the best answer to the problem. In one of our previous papers (Allik et al., 2020b), we already tried to correct rankings by taking into account in how many research areas each country/territory has failed to exceeded the entrance thresholds of the *ESI*. Although the spurious rankings were diminished, the improvement was less spectacular compared with the MDS of the disciplinary profiles used in this study. Additional studies are needed to establish what the best formula would be taking missing research fields into account.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

Adams, J., and King, C. (2010). *Global Research Report: Russia. Research and Collaboration in The New Geography of Science.* Leeds: Evidence.

Aksnes, D. W., Langfeldt, L., and Wouters, P. (2019). Citations, citation indicators, and research quality: an overview of basic concepts and theories. *Sage Open* 9:2158244019829575. doi: 10.1177/2158244019829575

Albarran, P., Perianes-Rodriguez, A., and Ruiz-Castillo, J. (2015). Differences in citation impact across countries. *J. Assoc. Inform. Sci. Technol.* 66, 512–525. doi: 10.1002/asi.23219

Allik, J. (2013). Factors affecting bibliometric indicators of scientific quality. *Trames J. Human. Soc. Sci.* 17, 199–214. doi: 10.3176/tr.2013.3.01

Allik, J., Lauk, K., and Realo, A. (2020a). Factors predicting the scientific wealth of nations. *Cross. Cult. Res.* 54, 364–397. doi: 10.1177/1069397120910982

Allik, J., Lauk, K., and Realo, A. (2020b). Indicators of the scientific impact of nations revisited. *Trames J. Human. Soc. Sci.* 24, 231–250. doi: 10.3176/tr.2020.2.07

Almeida, J. A. S., Pais, A., and Formosinho, S. J. (2009). Science indicators and science patterns in Europe. *J. Informetr.* 3, 134–142. doi: 10.1016/j.joi.2009.01.001

Bongioanni, I., Daraio, C., Moed, H. F., and Ruocco, G. (2015). *Comparing the Disciplinary Profiles of National and Regional Research Systems by Extensive and Intensive Measures*. Avaliable online at: https://pdfs.semanticscholar.org/4406/daed0bde9f3367b1596b97d1fcc3aac360c4.pdf

Bongioanni, I., Daraio, C., and Ruocco, G. (2014). A quantitative measure to compare the disciplinary profiles of research systems and their evolution over time. *J. Informetr.* 8, 710–727. doi: 10.1016/j.joi.2014.06.006

Borg, I., and Groenen, P. J. F. (2005). *Modern Multidimensional Scaling: Theory and Applications, 2nd Edn*. New York, NY: Springer.

Bornmann, L., Wagner, C., and Leydesdorff, L. (2015). BRICS countries and scientific excellence: a bibliometric analysis of most frequently cited papers. *J. Assoc. Inform. Sci. Technol.* 66, 1507–1513. doi: 10.1002/asi.23333

Carley, S., Porter, A. L., Rafols, I., and Leydesdorff, L. (2017). Visualization of disciplinary profiles: enhanced science overlay maps. *J. Data Inform. Sci.* 2, 68–111. doi: 10.1515/jdis-2017-0015

Cimini, G., Gabrielli, A., and Labini, F. S. (2014). The scientific competitiveness of nations. *PLoS ONE* 9:e113470. doi: 10.1371/journal.pone.0113470

Confraria, H., Godinho, M. M., and Wang, L. L. (2017). Determinants of citation impact: a comparative analysis of the Global South versus the Global North. *Res. Policy* 46, 265–279. doi: 10.1016/j.respol.2016.11.004

Daraio, C., Fabbri, F., Gavazzi, G., Izzo, M. G., Leuzzi, L., Quaglia, G., and Ruocco, G. (2018). Assessing the interdependencies between scientific disciplinary profiles. *Scientometrics* 116, 1785–1803. doi: 10.1007/s11192-018-2816-5

Erfanmanesh, M., Tahira, M., and Abrizah, A. (2017). The publication success of 102 Nations in scopus and the performance of their scopus-indexed journals. *Publish. Res. Q.* 33, 421–432. doi: 10.1007/s12109-017-9540-5

Garfield, E. (1955). Citation indexes for science: a new dimension in documentation through association of ideas. *Science* 122, 108–111. doi: 10.1126/science.122.3159.108

Glänzel, W. (2000). Science in Scandinavia: a bibliometric approach. *Scientometrics* 48, 121–150. doi: 10.1023/A:1005640604267

Guskov, A., Kosyakov, D., and Selivanova, I. (2016). Scientometric research in Russia: impact of science policy changes. *Scientometrics* 107, 287–303. doi: 10.1007/s11192-016-1876-7

Hakonarson, H., Gulcher, J. R., and Stefánsson, K. (2003). deCODE genetics, Inc. *Pharmacogenomics* 4, 209–215. doi: 10.1517/phgs.4.2.209.22627

Halffman, W., and Leydesdorff, L. (2010). Is inequality among universities increasing? gini coefficients and the elusive rise of elite Universities. *Minerva* 48, 55–72. doi: 10.1007/s11024-010-9141-3

Harzing, A. W., and Giroud, A. (2014). The competitive advantage of nations: an application to academia. *J. Informetr.* 8, 29–42. doi: 10.1016/j.joi.2013.10.007

Horta, H. (2018). The declining scientific wealth of Hong Kong and Singapore. *Scientometrics* 117, 427–447. doi: 10.1007/s11192-018-2845-0

Ivanova, M. (2007). Designing the United Nations environment programme: a story of compromise and confrontation. *Int. Environ. Agree. Politic. Law Econom.* 7, 337–361. doi: 10.1007/s10784-007-9052-4

Jurajda, S., Kozubek, S., Munich, D., and Skoda, S. (2017). Scientific publication performance in post-communist countries: still lagging far behind. *Scientometrics* 112, 315–328. doi: 10.1007/s11192-017-2389-8

Karlsson, S., and Persson, O. (2012). *The Swedish Production of Highly Cited Papers*. Stockholm: Vetenskapsrådet.

King, D. A. (2004). The scientific impact of nations. *Nature* 430, 311–316. doi: 10.1038/430311a

Kozlowski, J., Radosevic, S., and Ircha, D. (1999). History matters: the inherited disciplinary structure of the post-communist science in countries of Central and Eastern Europe and its restructuring. *Scientometrics* 45, 137–166. doi: 10.1007/BF02458473

Lauk, K., and Allik, J. (2018). A puzzle of Estonian science: how to explain unexpected rise of the scientific impact. *Trames J. Human. Soc. Sci.* 22, 329–344. doi: 10.3176/tr.2018.4.01

Leydesdorff, L., and Wagner, C. (2009). Is the United States losing ground in science? A global perspective on the world science system. *Scientometrics* 78, 23–36. doi: 10.1007/s11192-008-1830-4

Leydesdorff, L., Wagner, C. S., and Bornmann, L. (2014). The European Union, China, and the United States in the top-1% and top-10% layers of most-frequently cited publications: competition and collaborations. *J. Informetr.* 8, 606–617. doi: 10.1016/j.joi.2014.05.002

Li, N. (2017). Evolutionary patterns of national disciplinary profiles in research: 1996-2015. *Scientometrics* 111, 493–520. doi: 10.1007/s11192-017-2259-4

Lorca, P., and de Andrés, J. (2019). The importance of cultural factors in R&D intensity. *Cross Cult. Res.* 53, 483–507. doi: 10.1177/1069397118813546

MacRoberts, M. H., and MacRoberts, B. R. (2018). The mismeasure of science: citation analysis. *J. Assoc. Inform. Sci. Technol.* 69, 474–482. doi: 10.1002/asi.23970

Mair, P., Borg, I., and Rusch, T. (2016). Goodness-of-fit assessment in multidimensional scaling and unfolding. *Multivar. Behav. Res.* 51, 772–789. doi: 10.1080/00273171.2016.1235966

Markusova, V. A., Jansz, M., Libkind, A. N., Libkind, I., and Varshavsky, A. (2009). Trends in Russian research output in post-Soviet era. *Scientometrics* 79, 249–260. doi: 10.1007/s11192-009-0416-0

May, R. M. (1997). The scientific wealth of nations. *Science* 275, 793–796. doi: 10.1126/science.275.5301.793

Monge-Najera, J., and Ho, Y. S. (2015). Bibliometry of Panama publications in the science citation index expanded: publication type, language, fields, authors and institutions. *Rev. Biol. Trop.* 63, 1255–1266. doi: 10.15517/rbt.v63i4.21112

Öquist, G., and Benner, M. (2015). "Why are some nations more successful than others in research impact? a comparison between Denmark and Sweden," in *Incentives and Performance: Governance of Research Organizations,* eds I. M. Welpe, J. Wollersheim, S. Ringelhan, and M. Osterloh (Cham: Springer International Publishing), 241–257.

Pinto, T., and Teixeira, A. A. C. (2020). The impact of research output on economic growth by fields of science: a dynamic panel data analysis, 1980–2016. *Scientometrics* 123, 945–978. doi: 10.1007/s11192-020-03 419-3

Prathap, G. (2017). Scientific wealth and inequality within nations. *Scientometrics* 113, 923–928. doi: 10.1007/s11192-017-2511-y

Radosevic, S., and Yoruk, E. (2014). Are there global shifts in the world science base? Analysing the catching up and falling behind of world regions. *Scientometrics* 101, 1897–1924. doi: 10.1007/s11192-014-1344-1

Rousseau, S., and Rousseau, R. (1998). The scientific wealth of European nations: taking effectiveness into account. *Scientometrics* 42, 75–87. doi: 10.1007/BF02465013

Rubinoff, I., and Leigh, E. G. (1990). Dealing with diversity: the Smithsonian tropical research Institute and tropical biology. *Trends Ecol. Evolut.* 5, 115–118. doi: 10.1016/0169-5347(90)90165-A

Shashnov, S., and Kotsemir, M. (2018). Research landscape of the BRICS countries: current trends in research output, thematic structures of publications, and the relative influence of partners. *Scientometrics* 117, 1115–1155. doi: 10.1007/s11192-018-2883-7

Thelwall, M., and Levitt, J. M. (2018). National scientific performance evolution patterns: Retrenchment, successful expansion, or overextension. *J. Assoc. Inform. Sci. Technol.* 69, 720–727. doi: 10.1002/asi.23969

Tregubova, N. D., Fabrykant, M., and Marchenko, A. (2017). Countries versus disciplines: comparative analysis of post-soviet transformations in academic publications from Belarus, Russia and Ukraine. *Comparat. Sociol.* 16, 147–177. doi: 10.1163/15691330-12341414

van Leeuwen, T. N., Visser, M. S., Moed, H. F., Nederhof, T. J., and van Raan, A. F. J. (2003). Holy Grail of science policy: exploring and combining bibliometric tools in search of scientific excellence. *Scientometrics* 57, 257–280. doi: 10.1023/A:1024141819302

Wagner, C. S., Whetsell, T., Baas, J., and Jonkers, K. (2018). Openness and impact of leading scientific countries. *Front. Res. Metr. Anal.* 3:10. doi: 10.3389/frma.2018.00010

Waltman, L. (2016). A review of the literature on citation impact indicators. *J. Informetr.* 10, 365–391. doi: 10.1016/j.joi.2016.02.007

Yang, L. Y., Yue, T., Ding, J. L., and Han, T. (2012). A comparison of disciplinary structure in science between the G7 and the BRIC countries by bibliometric methods. *Scientometrics* 93, 497–516. doi: 10.1007/s11192-012-0695-8

Yue, T., Yang, L. Y., Ahlgren, P., Ding, J. L., Shi, S. Q., and Frietsch, R. (2018). A comparison of citation disciplinary structure in science between the G7 countries and the BRICS countries. *J. Data Inform. Sci.* 3, 14–31. doi: 10.2478/jdis-2018-0012

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Research Metrics and Analytics

SYSTEMATIC REVIEW
published: 28 October 2020
doi: 10.3389/frma.2020.551147

Bibliometric and Visualized Analysis of China's Smart Grid Research 2008–2018

Cheng Wang[1,2], Tao Lv[1]* and Xu Deng[1]

[1] School of Management, China University of Mining and Technology, Xuzhou, China, [2] School of Foreign Languages, Huaiyin Normal University, Huai'an, China

Smart grid (SG) offers great advantages in renewable energy integration and has become a popular trend of modern power development recently; meanwhile China is the second most prolific country using SG. Hence the purpose of this study is to get access to the research status, development, and trends of SG in China based on the 3,558 published papers obtained from the WOS core library and application of the bibliometric method and visualization analysis software VOSviewer and alluvial diagrams. The results consequently demonstrate some valuable insights. Firstly, the volume of publications in China's SG is on the rise, and the cooperation between countries and institutions is getting closer. Besides, the research hotspots have obvious interdisciplinary characteristics. Taking into consideration the impact of the information and communication field on SG, the major current research hotspots include wireless sensor network (WSN), internet of things (IoT), smart meter, big data, and security. Taking into consideration the impact of SG on traditional power systems, the main hotspots cover demand response, micro-grid, distributed generation, and electric vehicle (EV). Furthermore, China's SG research shows a trend from a single theme to diversified development. The research themes during 2010–2018 have deepened with most studies focusing on the traditional power system. The findings of this paper provide some enlightenment on China's SG research, which can present scholars with an overview of the macro perspective, help them understand the latest development of the SG field in China and offer useful guidance for future research in this subject as well.

Keywords: smart grid (SG), bibliometrics, co-authorship, co-occurrence, VOSviewer, alluvial diagram, China

OPEN ACCESS

Edited by:
Zaida Chinchilla-Rodríguez, Consejo Superior de Investigaciones Científicas (CSIC), Spain

Reviewed by:
Carlos Mulet Forteza, University of the Balearic Islands, Spain
Carlos Olmeda-Gómez, Universidad Carlos III de Madrid, Spain

****Correspondence:***
Tao Lv
taocumt@cumt.edu.cn

Specialty section:
This article was submitted to Research Assessment, a section of the journal Frontiers in Research Metrics and Analytics

Received: *14 May 2020*
Accepted: *18 August 2020*
Published: *28 October 2020*

Citation:
Wang C, Lv T and Deng X (2020) Bibliometric and Visualized Analysis of China's Smart Grid Research 2008–2018. Front. Res. Metr. Anal. 5:551147. doi: 10.3389/frma.2020.551147

INTRODUCTION

The integration of large-scale renewable energy in power systems is inevitable, but large-scale renewable energy is intermittent and variable, which has a great impact on the power grid. Smart grid (SG) with the ability of rapid response and self-repair has become an important solution to the above problems, as SG can promote the use of clean energy, improve the efficiency of power generation and energy utilization, improve the transmission efficiency of power grids, and improve the energy efficiency of terminals (Yuan and Hu, 2011). Consequently, SG has become a recent developmental trend for the world power grid, which can improve the efficiency and reliability of the power grid, reduce peak demand (El-Hawary, 2017), realize automatic control, and self-repair (Gungor et al., 2011).

SG is not only a single technology but also a series of new technologies and institutional innovations that can make the grid more efficient, cleaner, and smarter. Yu et al. define SG in China as "*an integration of renewable energy, new materials, advanced equipment, information technology, control technology, and energy storage technology, which can realize digital management, intelligent decision making, and interactive transactions of electricity generation, transmission, deployment, usage, and storage*" (Yu et al., 2012). SG has several new functionalities: self-healing, motivating the consumer, resisting attack, increasing power quality, accommodating all generation and storage options, enabling electrical markets, optimizing assets, and operating efficiently (Baumeister, 2010). In recent years, SG has attracted more and more attention, and the number of related papers published has increased. So far, more than 20,000 related papers have been published in different journals.

On the basis of qualitative analysis, many scholars have published reviews on SG. Technology development is the primary condition for the development of SG, which mainly involves wireless sensor technology (Mahmood et al., 2015), communication technology (Usman and Shami, 2013; Kabalci, 2016), artificial intelligence technology (Zhang et al., 2018), big data technology (Tu et al., 2017), and internet of things technology (Hossain et al., 2019). Institutional innovations are another important area. Along with technological innovation, developed and developing countries face similar challenges in the SG area, and national incentives and national energy resources limit the development of SG (Ponce-Jara et al., 2017). The pioneers of developing countries in SG are China, India, and Brazil (Fadaeenejad et al., 2014). Yu et al. gave a comprehensive overview of China's SG development, obstacles and barriers, and policy prospects (Yu et al., 2012). Yuan et al. analyzed the policy, pilot projects, achievements, and barriers of developing SG in China, and found that the lack of national strategy and the current industrial structure of the power industry were obstacles to its development (Yuan et al., 2014).

Although these existing reviews are beneficial for scholars to understand the SG development, they have merely taken a qualitative approach to review the content and subject matter of published literature. More knowledge, therefore, can be gained by quantitatively analyzing the existing literature and exploring and tracking the evolution of a large number of published work.

At present, a large number of scholars have used bibliometric methods to quantitatively visualize the landscape and the evolution of various scientific research fields (e.g., Montoya et al., 2014; Guo et al., 2019; Merigó et al., 2019; Shi and Liu, 2019). These studies are of great help to our quantitative analysis of SG research in China. For example, 10,938 journal articles, and 144 books on SG from ScienceDirect from 2008 to 2015 were reviewed and discussed the features, functionalities, and characteristics of SG (Tuballa and Abundo, 2016). However, this paper only briefly analyzes the annual publication quantity and technical classification of literature, and subjectively selects some literature for analysis. Coincidentally, Hossain et al. searched relevant literature on SG from databases such as Elsevier, Springer, Taylor & Francis, and Wiley for analysis, and determined the role of SG in renewable energy (Hossain et al., 2016). This paper also searched the literature in SG from different databases and only selected part of the literature for categorization and analysis.

The above two works of literature, analyzed from a literature perspective, play an important role in the SG field. However, there are still some defects, mainly in the following aspects: First of all, these reviews only selected a few representative papers. Secondly, the author selects the literature subjectively according to his own experience. Thirdly, existing review articles rarely cover the research hotspots, cooperative networks, and development trends of SG. However, there are currently still too few bibliometric papers in the SG field.

Quantitative visualization analysis of the SG field in China is very useful because it can supplement and verify qualitative reviews (Zhu et al., 2019). Compared with qualitative reviews, quantitative reviews can elucidate the status quo and development within the SG field in China from a macro perspective and provide an objective and intuitive overview. Besides, with the help of visualization tools, the landscape, and evolutionary patterns of China's SG can be more intuitively displayed. Therefore, it is important and timely to conduct a quantitative review of China's SG.

In this review, we attempt to quantify the landscape and development trajectory of SG research in China and discover the current research frontiers. We reviewed 11 years of published research (2008–2018) from the WOS Core library and used the visualization tool VOSviewer to detect, quantify, and visualize the current status and evolution of SG research in China.

Our bibliometric review has the following possible contributions. Firstly, it provides a new way to discover partnerships through co-authorship and is able to show it visually. Secondly, the evolution of the frontier of SG research in China is quantitatively tracked by stages. Thirdly, this review provides an overview of SG research in China from a macro perspective. Therefore, this review can provide scholars with a systematic understanding of the status quo, research frontiers, and future trends of SG research in China, and thus promote the development of SG research in the future.

The present study aimed to answer the following research questions of China's SG:

(1) What are the publication, citation status, and trend of smart grid literature? This will help researchers identify SG trends and predict future patterns in the field.
(2) Which are the most influential institutions in China's SG field and what are the differences in their research hotspots? How is China's cooperation in the SG field? This will help the researcher identify the subject and potential research collaborators.
(3) What are the most influential papers and journals? This will help researchers to consider which journals to choose to publish their manuscripts in within the SG field, potentially affecting future citations of their literature. On the other hand, most influential papers will help researchers and practitioners gain access to the literature that needs the most attention in the field. It will be beneficial for researchers to find research directions and methods.

(4) What are China's research hotspots in the SG field? How do research hotspots evolve? This will help researchers to understand the research direction and development trend of SG in China from a global perspective, and point out the direction for future research.

To solve the above research questions, this paper is organized as follows: In section Methodology, we will introduce the research methods, including the database selection, retrieval strategy, bibliometric methods, and data cleaning. A general overview of the SG field in China will be presented in section Basic Features in SG of China, including annual changes in publications and citations, the most published Chinese institutions and influences, the most published journals, and key articles ranked by citations. Section Visualization Analysis displays a visual presentation of keyword co-occurrence and discovers research hotspots and their evolution process in three stages. In the last section, we will present the conclusions and limitations.

METHODOLOGY

This paper provides a systematic review of scientometric analysis in SG of China. **Figure 1** shows the research design of this study.

Data Collection and Processing

Before the bibliometric analysis, it is necessary to establish a data set containing citation information. Currently, the citation databases include Scopus, ISI Web of Science (WOS), and Google Scholar. Olawumi et al. compared the advantages and disadvantages of the three databases (Olawumi et al., 2017). This paper finally selects the Science Citation Index Science (SCIE), Conference Proceedings Citation Index-Science (CPCI-S) and Social Sciences Citation Index (SSCI), which are in the core library of the WOS database. The WOS records were chosen as they contain the most comprehensive and influential journals, which have scientific robustness. Therefore, they have become the choice of many scholars for literature measurement (Rahman et al., 2017; Olawumi and Chan, 2018; Yu and He, 2020).

The search strategy adopted to retrieve the paper data on China's SG was as follows:

TS = ("smart grid" OR "smartgrid*"). Document Type: Article OR Proceedings Paper OR Review. Timespan = 2008–2018. Databases: Science Citation Index Expanded (SCIE), Conference Proceedings Citation Index – Science (CPCI-S), Social Sciences Citation Index (SSCI). Country/Region: PEOPLES R CHINA. Search time: 2019-11-14 (SCIE and CPCI-S), 2020-03-20 (SSCI). TS refers to topic research in Web of Science.*

Before we can analyze the data, we must clean the sample we received from the WOS (Mulet-Forteza et al., 2019). In the co-authorship analysis, countries that appear in different names but choose to be recognized as a single country have been unified, such as "Scotland" and "England" were united as the "United Kingdom." "Taiwan" and "Hong Kong" were united as "People R China." Also, the organization's name is unified. For example, "NCEPU" and "N CHINA ELECT POWER UNIV" are uniformly modified to "North China Elect Power University." Furthermore, subordinate agencies of the State Grid Corporation of China will be integrated into the State Grid Corporation of China. For example, the names of "Guangdong POWER Grid Corporation," "Jiangsu POWER Grid Corporation," and "CHINA POWER SCI RES INS" will be replaced by "State Grid Corporation of China."

Finally, in the co-occurrence of author-keywords, First of all, for the lack of author keywords, the necessary supplements were made against the original literature. Besides, different expressions of keywords have been integrated, such as unifying "smart grids," "smart grid," and "smart-grids" into "smart grid"; "V2G" and "Vehicle-to-grid" have been unified into "Vehicle-to-grid (V2G)." "Demand response," "DR" unified into "demand response (DR)."

Scientometric Analysis Methods

In this article, we use a variety of bibliometric methods to analyze the acquired data set. Firstly, the number of publications (TP) is used to detect the quantified productivity (Ding et al., 2014), meanwhile, the number of citations (TC) is used to measure influence (Goran, 2010). "Citations per document" (TC/TP) shows the average impact per paper. Another bibliographic method is h-index, which means h number of papers published in a journal, or by an organization, have at least h citations (Hirsch, 2005). Some other common methods include the most productive institutions, journals, the most cited articles, average number of authors per article (AN/TP), average number of references per article (RN/TC), and impact factor (IF) (Garfield, 1983).

Finally, we use VOSviewer software (see www.vosviewer.com) to visualize the graphical mapping of the bibliographic data. VOSviewer can construct bibliometric networks based on data from WOS, Scopus, Dimensions, and PubMed files, or reference manager files (i.e., RIS, EndNote, and RefWorks files). It uses distance-based maps to construct a co-authorship map, co-occurrence map, citation, bibliographic coupling, and co-citation map (van Eck and Waltman, 2010). In this article, co-authorship was used for country and institution cooperation analysis. Since the early 1980s, co-authorship has been operating as a proxy for research cooperation (Subramanyam, 1983). Co-occurrence was used for author-keywords analysis (He, 1999).

There are two counting methods in VOSviewer: full counting and fractional counting. Perianes-Rodriguez et al. have compared the two methods of counting, they think that in many cases, the difference is relatively limited. There may not be a conclusion from the literature on network analysis to produce fundamental influence, especially when the conclusion is based on the analysis of the datasets. Using full counting or fractional counting, there is no essential difference between the results obtained (Perianes-Rodriguez et al., 2016). Therefore, full counting was chosen as the counting method in this paper.

BASIC FEATURES IN SG OF CHINA

We found that there were 20,195 research papers about SG around the world, among the highest were the United States with 4,457 and China with 3,558. China publishes 324 articles every year on average, so it is meaningful to conduct a global analysis

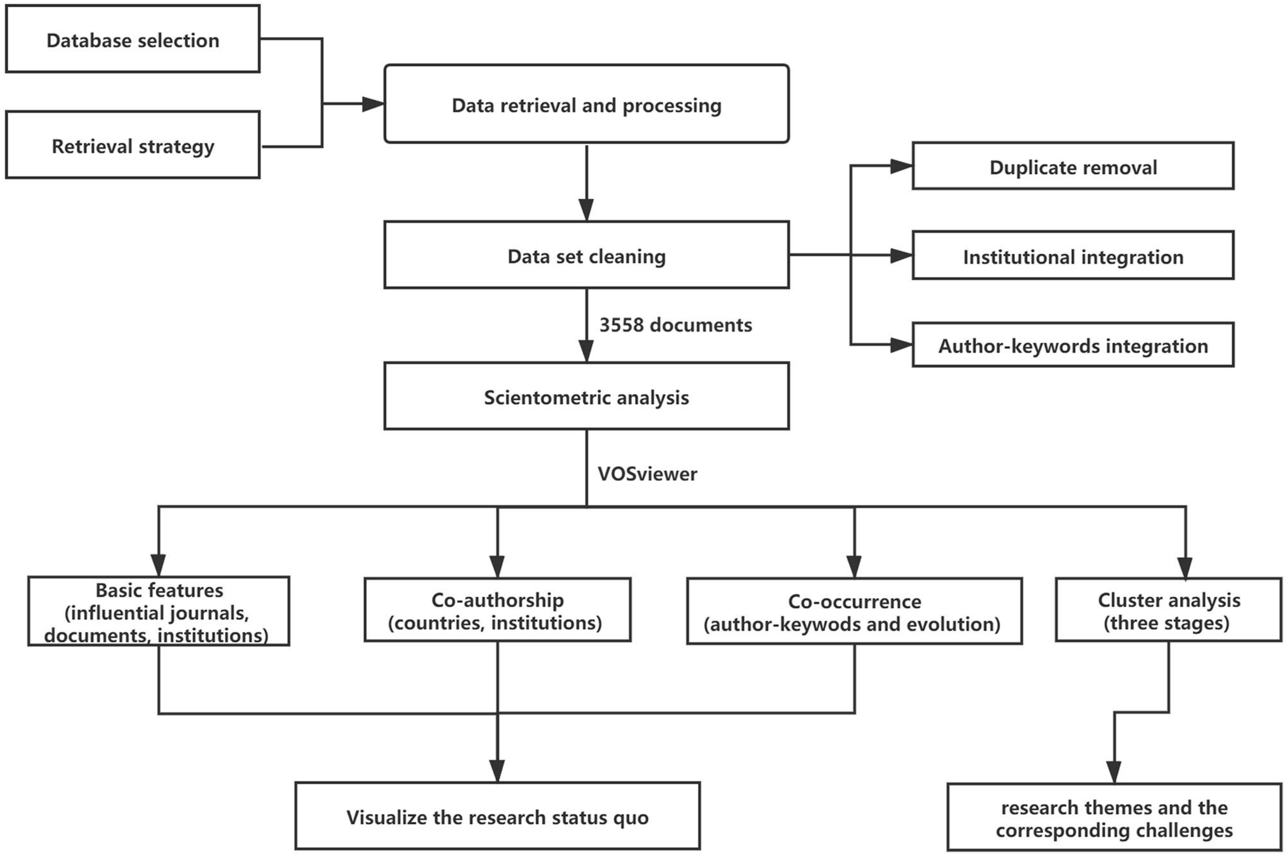

FIGURE 1 | Outline of research design.

on the SG field in China. The term "Chinese scholar" in this paper refers to a scholar from a Chinese institution where the author published the document.

Annual Publications and Growth Trend

As set out in **Table 1**, the papers published by Chinese scholars before 2009 were no more than 10 per year, but after the year 2010 the number of publications increased significantly. Then there was a small decline in 2015, followed by a slow increase and a steady trend in 2016. Therefore, the research literature in the SG field of China can be divided into three stages: the embryonic stage before 2009, the developmental stage from 2010 to 2014, and the stable stage from 2015 to 2018.

Figure 2 presents a comparative analysis of the publication numbers of Chinese and worldwide scholars. Global publication began to decline after peaking in 2016, but the number of publications by Chinese scholars continued to rise until 2018. The number of articles published by Chinese scholars has increased year by year, and more than 570 were published in 2018. Besides, since 2016, Chinese scholars have published more than 550 articles on SG every year.

TABLE 1 | Chinese scholars' publications characteristics on SG from 2008 to 2018.

Years	TP	AN	AN/TP	TC	TC/TP	RN	RN/TP
2008	1	4	4	0	0	5	5
2009	8	39	5	15	2	145	18
2010	91	337	4	759	8	1,068	12
2011	140	508	4	1,466	10	1,812	13
2012	288	1,080	4	4,832	17	4,934	17
2013	354	1,390	4	7,949	22	6,574	19
2014	488	1,975	4	5,376	11	8,847	18
2015	476	2,084	4	4,679	10	10,473	22
2016	579	2,439	4	4,379	8	12,998	22
2017	558	2,504	4	4,164	7	15,019	27
2018	575	2,640	5	2,290	4	17,791	31
Average	323.5	1363.6	4.1	3264.5	9.1	7242.4	18.5
Total	3,558	15,000	–	35,909	–	79,666	–

TP, total publication; AN, author number; TC, total citation; RN, reference number.

China's SG publication volume is closely related to policy promotion. For example, in 2010, the National Energy Administration (NEA) started to promote smart grid standardization work by the "*Notice on The Establishment of National Smart Grid Standardization Overall Work Promotion Group*"(National Energy Administration, 2010). In the same year, the State Grid Corporation of China issued the "*Strong Smart*

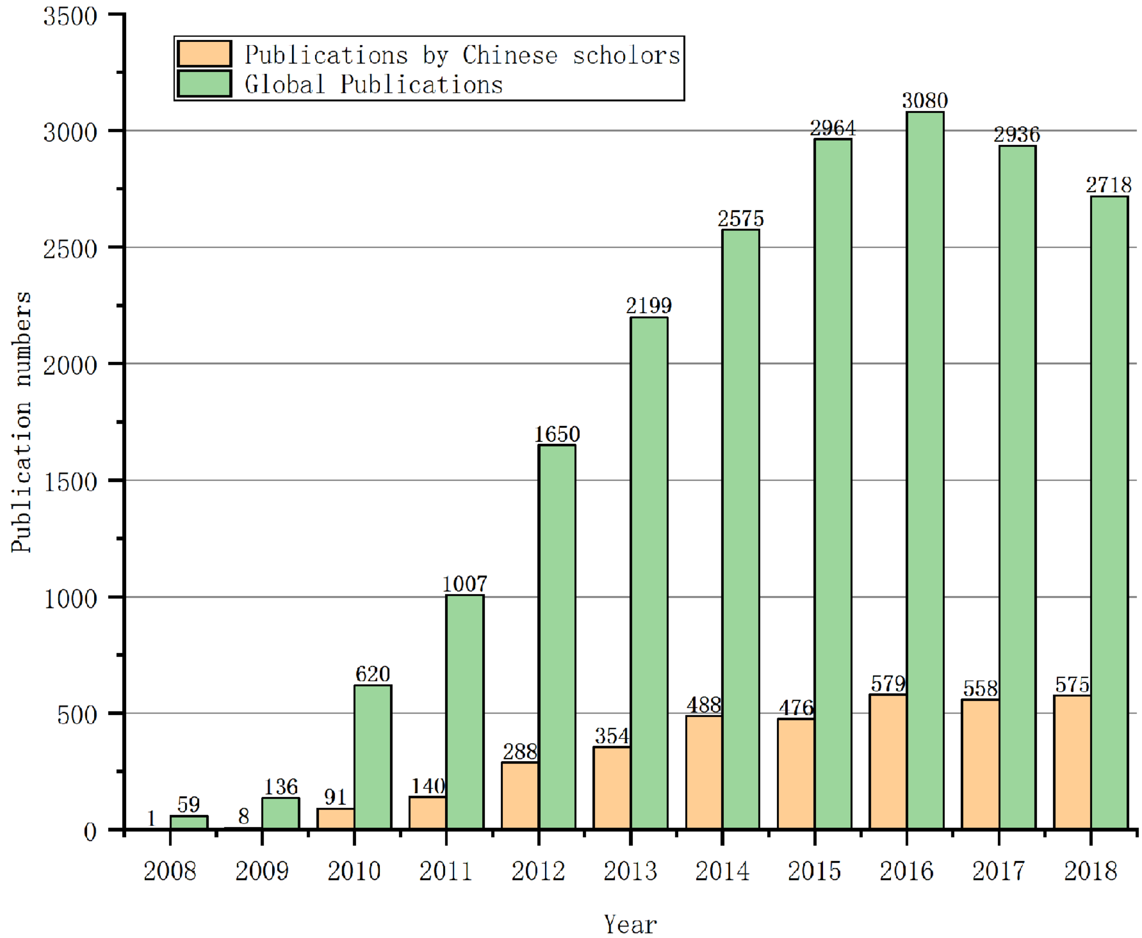

FIGURE 2 | Annual publications by Chinese scholars and worldwide scholars in the field of SG.

Grid Technical Standard System Plan," which officially carried out smart grid construction at the national level (Wang, 2010). Since 2010, relevant research began to greatly increase. Similarly, in 2015 the National Development and Reform Commission (NDRC) and NEA issued the "*Guidelines on Promoting the Development of Smart Grid*" (National Development and Reform Commission and National Energy Administration, 2015), which further promoted the development of China's SG. In the following year, China published more than 100 articles on SG. In 2016, NDRC and NEA released "*the 13th Five-Year Plan of Electric Power Development (2016–2020)*" (National Development Reform Commission and National Energy Administration, 2016) which proposed to optimize the power grid structure and further promote the construction of the smart power grid. From 2016 to 2018, the number of papers published in this period was maintained at over 550. In 2020, NDRC and the Ministry of Justice issued "*the Opinions on Accelerating the Establishment of Green Production and Consumption Regulations and Policy System*" (National Energy Administration Ministry of Justice, 2020) which proposed to increase policy support for distributed energy, smart grid, energy storage technology, and multi-energy complementary technology. It is expected that relevant policies will still be introduced in the 14th 5-Year Plan to promote the development of SG in China, and research literature on SG in China will continue to be maintained in the next few years.

Common bibliometric indicators are used in **Table 1**, for example, TP (total publication) refers to the total number of publications per year, AN (author number) represents the number of authors and can be used to describe the strength of collaboration between authors, TC (total citation) refers to the number of citations, usually used to describe the influence of the literature, RN (reference number) refers to the number of references cited, which is generally used to describe the basis of the literature research, and TC/TP represents the average annual citation amount of the literature which reflects the citation situation of the literature issued by Chinese scholars (Yu et al., 2020).

The number of citations per article by Chinese scholars in the SG field averaged <10 in 2008–2018 and reached the highest level of 22 in 2013. The cited frequency in China is relatively low, among which only two papers were cited more than 1,000 times,

TABLE 2 | Most productive Chinese institutions on SG.

Rank	Institution	TP	TC	TC/TP	h-index	Cited interval[a]				
						≥300	≥200	≥100	≥50	≥10
1	SGCC	644	1,652	2.6	17	0	1	1	2	38
2	NCEPU	348	1,653	4.8	19	0	1	1	4	39
3	THU	177	3,810	21.5	28	2	1	2	15	37
4	CAS	152	4,535	29.8	29	3	2	2	8	30
5	SJTU	149	1,298	8.7	20	0	0	2	3	25
6	ZJU	148	3,045	20.6	26	1	2	6	7	32
7	HKU	108	2,409	22.3	26	0	3	5	11	63
8	SEU	103	545	5.3	12	0	0	1	1	12
9	XJTU	95	1,256	13.2	17	0	1	0	5	25
10	HUST	86	1,850	21.5	18	1	0	1	5	19

TP, total publication; TC, total citation. cited interval refers to the annual total number of TC references interval, [a]the interval ≥200 in cited interval means 200≦cited interval <300, and all others are similar intervals. Institution abbreviation: SGCC, State Grid Corporation of China; NCEPU, North China Elect Power University; THU, Tsinghua University; CAS, Chinese Academy of Sciences; SJTU, Shanghai Jiaotong University; ZJU, Zhejiang University; HKU, University of Hong Kong; SEU, Southeast University; XJTU, Xi'an Jiao Tong University; HUST, Huazhong University of Science and Technology.

one paper was cited more than 500 times, and 1,722 articles have not been cited since publication.

The Most Productive Chinese Institutions

The top 10 institutions have published a total of 2010 papers, accounting for 56.5%. These 10 institutions are China's main research institutions in the SG field. **Table 2** shows some characteristics of these 10 institutions, such as TP, TC, and TC/TP. SGCC (State Grid Corporation of China) ranks first with 644 articles and accounts for 18.1%. Then there are NCEPU (North China Electric Power University), THU (Tsinghua University), CAS (Chinese Academy of Sciences), and SJTU (Shanghai Jiaotong University). There are two institutions whose TP quantity is >300, six institutions between 100 and 200. Both institutions ranked 9 and 10 have a TP quantity <100.

In terms of TC/TP, the highest is CAS, with an average of 29.8 citations per document. The second to the fifth most cited institutions are HKU (University of Hong Kong) (22.3), THU (21.5), HUST (Huazhong University of Science and Technology) (21.5), and ZJU (Zhejiang University) (16.9). It should be pointed out that SGCC is the institution with the most articles published, but its TC/TP ranked only 10th.

H-index gives an estimate of the importance, significance, and the broad impact of a scientist's cumulative research contributions (Hirsch, 2005). The CAS ranks first in the h-index with the value of 29, followed by THU (28), ZJU, and HKU with an h-index value of 26. There is a big difference between TP and h-index in some institutions, such as HUST which was tenth in TP, which ranks 7th in the h-index, and NCEPU, which ranks 2nd in TP but 6th in h-index.

Table 2 also gives new indicators: cited interval. Cited interval ≥300 refers to the number of literature published by an institution whose citation quantity is ≥300. In this interval, only three papers had been published by the CAS. Besides, one institution had published two articles, two institutions published only one article in this interval, and six institutions did not publish any literature. In institutions with more than 10 citations, HKU (82) ranks first, followed by THU (57), ZJU (48), CAS (45), and NCEPU (45). Cited interval can reflect the academic influence of relevant institutions.

In terms of the frequency of institutional citation, the literature published by the CAS in the SG field was cited the most, more than 4,500 times, with an average of 29.8 citations per article. According to the citation interval, only seven papers were cited more than 300 times in the top 10 institutions, 11 papers were cited more than 200 times, and most of the papers were cited <10 times.

The State Grid Corporation of China (SGCC) published the most literature, but the Chinese Academy of Science (CAS) had the highest h-index, cited interval >300 had more literature than the other institutions, indicating that CAS was the most influential institution in the SG field in China.

The Most Productive Journals

Overall, 3,558 articles by Chinese scholars were published in 1,268 journals, with an average of 2.8 articles published in each journal. As far as productive journals are concerned, **Table 3** shows the top 10 most productive journals published by Chinese scholars on SG and other relevant information. The top 10 journals had published 478 Chinese scholars' papers in the SG field, accounting for only 13.8%. Nevertheless, 41% of journals published only one article and 14% of journals published two articles on the above.

Journal publications are relatively scattered. *IEEE Transactions on Smart Grid*, have published only 142 papers in the past 11 years. Among the top 10 journals, eight have an IF (2008) index >4. The largest IF (2008) is *Renewable & Sustainable Energy Reviews* with a value of 10.556, followed by *IEEE Transactions on Smart Grid* (10.486), and Applied Energy (8.426).

On the other hand, the largest TC/TP journal is *IEEE Transactions on Industrial Electronics*. The average number of

TABLE 3 | The top 10 most productive journals on SG by Chinese scholars.

Rank	Journal	IF (2018)	IF (5 years)	TP	TC	TC/TP	Cited interval				h	The most cited article	Times cited
							≥200	≥100	≥50	≥20			
1	IEEE T SMART GRID	10.486	10.607	142	5,795	40.8	4	5	29	44	121	Li et al., 2010	418
2	ENERGIES	2.707	2.99	71	524	7.4	0	0	0	7	64	Yu et al., 2012	39
3	IEEE ACCESS	4.098	4.54	54	571	10.6	0	0	2	7	56	Yu et al., 2017	95
4	IEEE T IND INFORM	7.377	8.423	53	1,893	35.7	2	3	4	13	100	Su et al., 2012	349
5	APPL ENERG	8.426	8.558	34	1,147	33.7	0	2	7	7	162	Wang et al., 2015	126
6	INT J ELEC POWER	4.418	4.262	30	666	22.2	1	0	2	4	100	Tan et al., 2013	261
7	RENEW SUST ENERG REV	10.556	11.239	30	1,020	34.0	0	3	3	9	222	Zhou et al., 2016	156
8	IEEE T POWER SYST	6.807	8.143	29	931	32.1	0	1	5	13	221	Zhong et al., 2013	141
9	IET GENER TRANSM DIS	3.229	3.432	25	232	9.3	0	0	0	3	94	Xiao et al., 2012	28
10	IEEE T IND ELECTRON	7.503	8.459	24	1,168	48.7	0	5	3	7	236	Strasser et al., 2015	164

IF (impact factor), the data were from the Journal Citation Reports database; TP, total publication; TC, total citation; h, h-index. Abbreviation of the journal name: IEEE T SMART GRID, IEEE Transactions on Smart Grid; IEEE T IND INFORM, IEEE Transactions on Industrial Informatics; APPL ENERG, Applied Energy; INT J ELEC POWER, International Journal Of Electrical Power & Energy Systems; IEEE T POWER SYST, IEEE Transactions On Power Systems; RENEW SUST ENERG REV, Renewable & Sustainable Energy Reviews; IEEE T IND ELECTRON, IEEE Transactions On Industrial Electronics, and IET GENER TRANSM DIS, IET Generation Transmission & Distribution.

citations for Chinese scholars published in this journal is 48.7, followed by *IEEE Transactions on Smart Grid* (40.8), and *IEEE Transactions on Industrial Informatics* (35.7). In terms of h-index, the highest value of *IEEE Transactions on Industrial Electronics* is 236, followed by *Renewable & Sustainable Energy Reviews* (222), and *IEEE Transactions on Power Systems* (221). **Table 3** also shows some indicators, such as the most cited literature in the journal and the number of citations, which can help scholars quickly find the important SG literature in the target journal.

Literature in the field of China's SG was mainly distributed in *Engineering*, *Electrical & Electronic*, *Computer Science*, *Energy & Fuels*, *Telecommunications*, *Automation & Control Systems*, and *Materials Science*, indicating that there are many interdisciplinary types of research on SG, and that authors of different disciplines are interested in SG.

IEEE Transactions on Smart Grid published the highest number of articles by Chinese scholars in 2008–2018, and it is also ranked second in IF and second in TC/TP, indicating that it is the most important journal in the SG field in terms of volume and influence.

The Most Cited Papers

Since 2008, Chinese scholars have published many influential papers in the SG field. **Table 4** lists the top 20 most cited papers in the SG field by Chinese scholars. Among them, the literature published by Pan et al. (2013) has been cited more than 1,650 times, and the other paper that was cited more than 1,000 times is Cheng and Chen (2012), both of which were completed by a single institution. Among the top 20 publications, multi-country cooperation accounted for 75% and multi-institution cooperation 80%. It can be seen that multi-country cooperation and multi-institution cooperation play an important role in high-impact research results in SG.

Table 4 also shows the cited references with high citation frequency in the SG field in China. Through the cited references with high citation frequency, research hotspots similar to the co-occurrence of keywords can be found. The top 20 cited literature can be divided into the following research topics: smart grid (rank 6, 10, 13), batteries and energy storage (rank 1, 2, 4, 5, 7, 12, 14, 19), EV (rank 8, 18), IoT (rank 9), demand response (rank 11, 15), smart grid communication technology (rank 16, 17, 20), and big data technology (rank 3).

TABLE 4 | The top 20 most cited papers on SG by Chinese scholars.

Rank	Most cited documents	TC	Citation/year	AN	RN	IN	CN
1	Pan et al., 2013	1,657	237	3	197	1	1
2	Cheng and Chen, 2012	1,300	163	2	217	1	1
3	Chen et al., 2014	852	142	3	155	3	2
4	Sun et al., 2013	475	68	11	56	4	2
5	Hu et al., 2013	437	63	3	278	1	1
6	Li et al., 2010	417	42	8	52	8	3
7	Wang et al., 2013	410	59	10	59	2	2
8	Su et al., 2012	349	44	4	91	2	2
9	Lin et al., 2017	290	97	6	163	5	2
10	Zhong et al., 2014	287	48	4	29	3	2
11	Deng et al., 2015	267	54	4	95	2	2
12	Rahimi-Eichi et al., 2013	262	38	4	45	3	2
13	Varaiya et al., 2011	261	29	3	44	3	3
14	Tan et al., 2013	260	38	3	52	2	1
15	Tsui and Chan, 2012	246	31	2	19	1	1
16	Zhang et al., 2012	243	31	6	10	2	2
17	Gao et al., 2012	237	30	5	203	4	2
18	Zhao et al., 2015	219	44	4	592	2	2
19	Wu et al., 2012	219	28	3	38	3	2
20	Liu et al., 2012	217	28	5	116	5	2

TC, total citation; AN, author number; RN, references number; IN, institution number; CN. country number.

The overall adoption of SG and dispatching synchronizer research are the main topics focusing on SG's impact on traditional power grids. For example, Li et al. divided SG into smart control center, smart transmission grid, and smart substation, and regarded it as an integrated system to discuss its characteristics and performance (Li et al., 2010). Zhong et al. focused on the self-synchronized synchronverters of the smart grid (Zhong et al., 2014). Varaiya et al. have designed a risk-and-speed dispatch mode to improve grid efficiency (Varaiya et al., 2011).

Batteries and energy storage account for the highest proportion of literature citations, which on the one hand reflects emphasis on energy storage in the SG field in China. Among them, three articles discussed room-temperature sodium-ion batteries (Pan et al., 2013; Sun et al., 2013; Wang et al., 2013), metal-air batteries were mentioned in one article (Cheng and Chen, 2012), LI4TI5O12-based electrodes for lithium-ion batteries were discussed in one article (Zhao et al., 2015), and one paper wrote about high-voltage lithium-ion batteries (Hu et al., 2013). In addition to the research of battery technology, the battery management system plays an important role in improving battery performance (Rahimi-Eichi et al., 2013). Energy storage systems (ESSS) realize comprehensive battery management from a wider range to improve the efficiency of battery use (Tan et al., 2013). EV are both users and suppliers in the smart grid, and their energy storage role is well-recognized. Wu et al. programmed a new game-theoretic model to understand the interactions among EV and aggregators in a vehicle-to grid (V2G) market (Wu et al., 2012). Su et al. focused on transportation electrification and introduced the current situation and prospect of electric vehicles in the field of industrial information systems (Su et al., 2012).

Network communication technology is the basis of smart grid operation, mainly including network security and privacy issues (Liu et al., 2012). The Machine-to-machine (M2M) Communications Paradigm (Zhang et al., 2012) and systematic Review of Communication/Networking Technologies in Smart Grid (Gao et al., 2012) discussed this.

IoT and big data play an irreplaceable role in the smart grid, Lin et al. reviewed the IoT in smart grid technology (Lin et al., 2017), similarly, Chen et al. reviewed the background and state-of-the-art of big data (Chen et al., 2014), These reviews play an important role in advancing the IoT and big data applications in SG.

Another important theme in China's SG field is demand response, which promotes load forecasting and management in smart grids (Tsui and Chan, 2012), power grid dispatch, and the electric market (Deng et al., 2015). The theme of highly cited literature in the SG field in China is similar to that of the co-occurrence of keywords, which can further reflect the research hotspots in the SG field in China.

VISUALIZATION ANALYSIS

Co-authorship Analysis

Co-authorship analysis has been widely used in the cooperative research of scientific research institutions and researchers. Although it is somewhat similar to the citation network (Garfield, 1979), co-authorship is a temporal and collegial relationship, which is fairer than the anonymity of citation (Liu et al., 2005).

Co-authorship Between Countries

Figure 3 shows the publication numbers of international cooperation and non-international cooperation, as well as the percentage of international collaborative publications. The number of international cooperation and non-international cooperation papers all showed a trend of growth and a larger percentage of international cooperation in 2009. This growth could possible be due to more published learning technology and experience from abroad at the initial stage of China's SG research. After 2010, the percentage of international cooperation fluctuated but has since been on the rise, which shows that the field is increasingly attracting the attention of more institutions.

The rate of international cooperation in the SG field is on the rise, but which countries have closer cooperation with China? **Figure 4** shows the network of countries with international cooperation with China. The United States is the country with the most cooperation, with 508 cooperative articles. Followed by 130 articles from Australia, 117 articles from the United Kingdom, and 109 articles from Canada. The larger the node is, the more papers published in cooperation between the country and China are, and the thicker the connection between the nodes, the closer the cooperation between the two countries is.

Figures 3, **4** show China's cooperation with other countries in the SG field. The current research also indicates that cooperation with foreign countries in the SG field is increasing. The average cooperation rate from 2008 to 2018 was 25.5%, and the cooperation rate reached 38% in 2018. The United States and China have the largest number of collaborative literature. Other countries that continue to work closely with China are Australia, the United Kingdom, and Canada. The overall increase of foreign cooperation rate indicates that Chinese scholars are paying more attention to cooperation with foreign countries in the SG field.

Co-authorship Between Institutions

As indicated in **Figure 5**, a total of 1,052 institutions have published literature in the SG field. The proportion of papers for inter-agency collaboration increased substantially in 2009, similar to the increase in country collaboration in 2009. In 2013, the number of inter-agency collaborations began to increase significantly, and the number of inter-agency collaboration papers in the following years showed an increasing trend. In terms of percentage, inter-agency collaboration was at 72% in 2015, fell to 65% in 2016, and then kept on rising, accounting for 78% in 2018.

To figure out the institutional cooperation in the SG field research by Chinese scholars, VOSviewer was used to make the institutional cooperation network as shown in **Figure 6**. The size of the circle in **Figure 6** is equal to the number of posts issued by the organization, and the thickness of the line represents the cooperative relationship between the organizations. For example, Fudan University and the University of Minnesota have the closest cooperation relationship (11 times), and then the State Grid Corporation of China and Zhejiang University

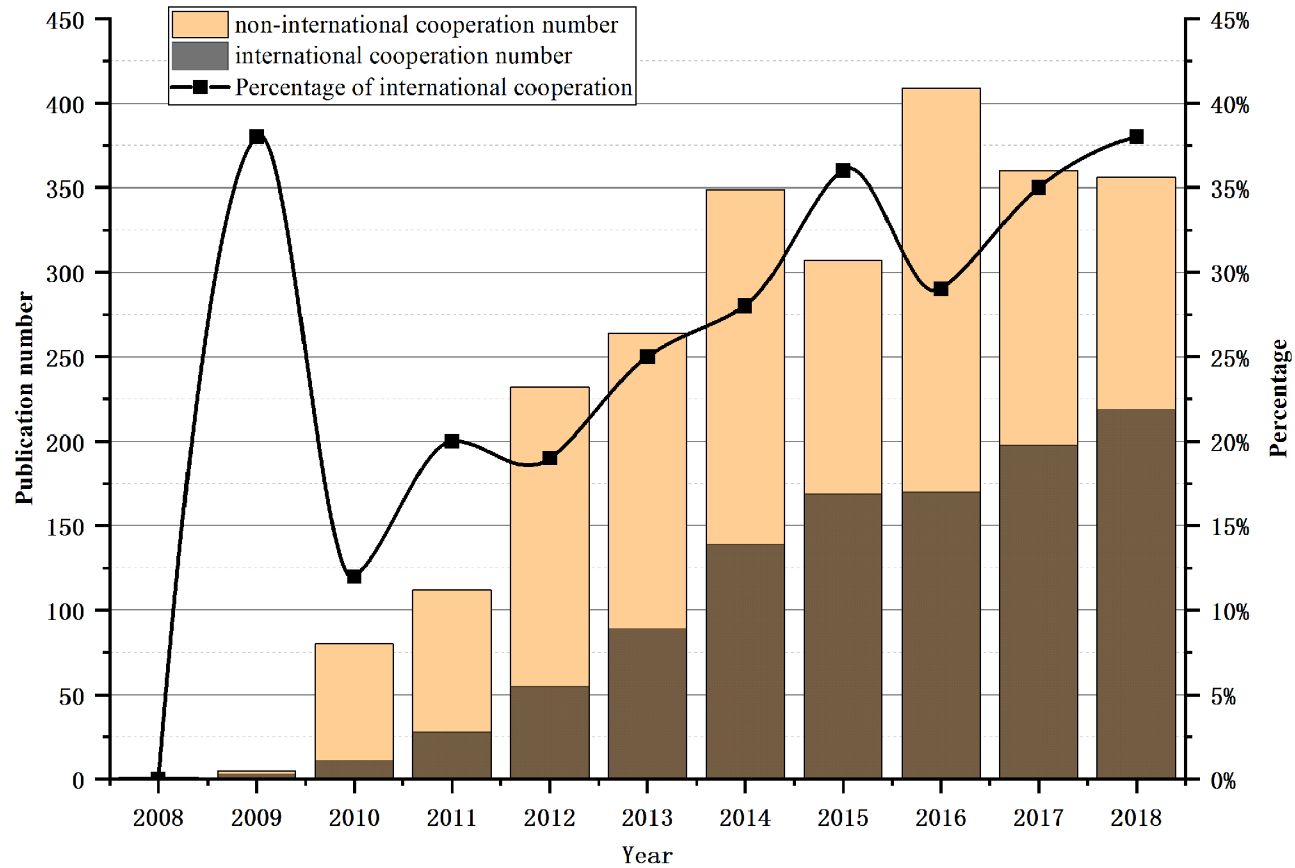

FIGURE 3 | Annual literature distribution of international and non-international cooperation.

(8 times), Guangdong University Technology, and University of Oslo (7 times).

Next we consider the cooperation clustering. The different colors in **Figure 6** represent the clustering of the VOSviewer to the institution, thereby revealing the collaboration between the institutions. Green is the cluster with the largest volume of publications, mainly including 10 institutions such as the State Grid Corporation of China, Zhejiang University, North China Electric Power University, and Tianjin University. The second is the purple clustering, which mainly represents Tsinghua University, Shanghai Jiao Tong University, Xi'an Jiao Tong University, Fudan University, and the University of Minnesota. Red clustering is far away, major representative institutions include the Chinese Academy of Science, Southeast University, Beijing Jiao Tong University, and the University of Electric Sci & Technology China. Blue clustering includes Huazhong University of SCI & Technology, Wuhan University, and the University of Oslo.

Co-occurrence Analysis

Co-occurrence analysis has been widely used in bibliometric analysis, the method of co-word analysis has been put forward since the 1980s (Callon et al., 1983). Keywords co-occurrence has been used at the forefront of the analysis of hot topics (Chen, 2017).

Author-Keywords Co-occurrence of Five Most Productive Institutions

Different institutions have different research emphases. **Figure 7** shows the keyword co-occurrence map of the five institutions with the most articles published in the SG field in China, which can be utilized to understand the research emphases of these five institutions. The hotspots of each institution are displayed as the depth of color in **Figures 7A–E**. The brighter the color, the more frequently the keyword appears. Demand response is the field with the brightest color in **Figures 7A–E**, which indicates that each institution regards the demand response as the research hotspot. The most frequent occurrence is State Grid Corporation of China (SGCC) (38 times), while the other institutions use this keyword more than 15 times. EV appeared in research hotspots of the State Grid Corporation of China (SGCC), North China Electric Power University (NCEPU), Tsinghua University (THU), and Shanghai Jiaotong University (SJTU). Big data is a common research hotspot of NCEPU, SGCC, and SJTU, while micro-grid is a research hotspot of NCEPU and SJTU.

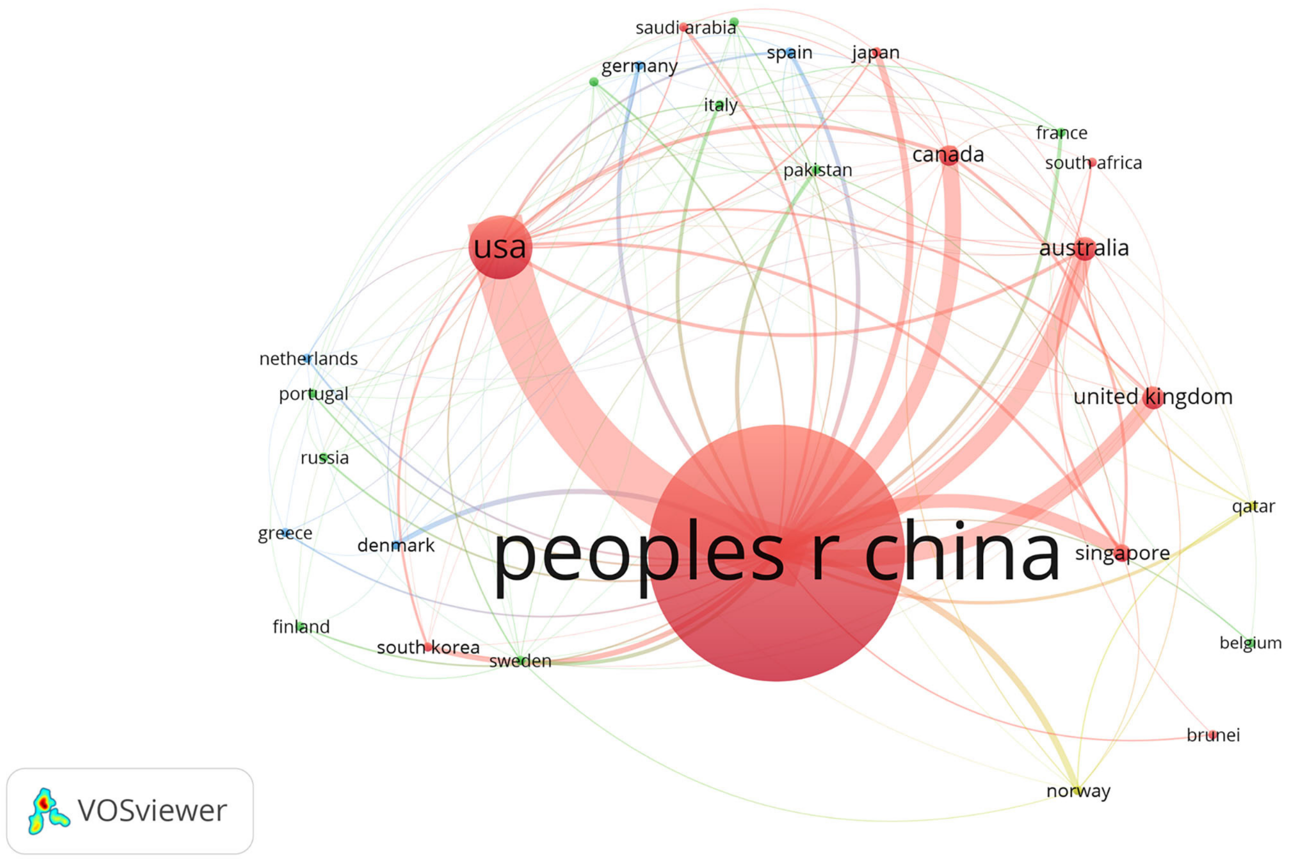

FIGURE 4 | Country co-authorship network in the field of the SG field. The larger the node is, the more papers published in cooperation between the country and China, and the thicker the connection between the nodes, the closer the cooperation between the two countries is. VOSviewer parameters: counting method: full counting, minimum occurrences 5, a total of 28 countries, normalization using LinLog/modularity method, layout parameters: attraction = 6, repulsion = 0.

In addition to the common research hotspot of demand response, different institutions have different research hotspots. For example, SGCC's research mainly focuses on big data, IoT, EV, smart substation, distribution network, cloud computing, WSN, distributed generation, micro-grid, renewable energy, IEC61850, and smart meter. SGCC, as China's largest network service provider, is committed to building a strong smart grid based on Ultra-High Voltage (UHV). This was proposed in 2009 by the smart grid development framework: one goal, two main lines, three stages, and four systems, five connotations, and six sections (Yu et al., 2012) including different aspects of smart grid construction. SGCC focused on IoT by smart electricity meters to improve data collection and analysis capabilities, the effectiveness of energy efficiency can be realized through big data and distributed systems. To improve energy efficiency, the use of electric vehicles and products such as vehicle-to-grid (V2G) that may affect the load on the grid needs to be considered. SGCC is the backbone of China's smart grid construction.

Except for SGCC, the other top four publishers are all scientific research institutions or universities, and their research hotspots are related to the characteristics of the institutions. For example, the hotspots of NCEPU (North China Electric Power University) involve various aspects of the power system, such as distributed generation, micro-grid, IoT, cloud computing, big data, and energy management information systems such as EV and energy storage. As a university focusing on power research, its University Council includes SGCC and other national key power companies and the China Electricity Council. The power system has always been its research topic. Its research hotspot includes SG technology foundation, such as IoT, power WSN, and many aspects of the power system.

Research hotspots for Tsinghua University mainly include EV, electricity market, cyber-physical systems (CPS), and smart meters. Among them, CPS is its unique research hotspot. Under the academic advantages of Tsinghua University, CPS achieves breakthroughs in remote control of distribution networks and reliable, safe, and efficient transmission and distribution of energy. CAS research mainly focuses on adaptive dynamic programming, optimal control, micro-grid, WSN, adaptive critic designs, security, neural network, and home energy management. CAS also pays attention to the research of power system planning and other related policies and algorithms. SJTU mainly studies renewable energy, WSN, micro-grid, IEC61850, batteries, energy management, and EV. SJTU pays more attention to renewable energy utilization, WSN, micro-grid, and IEC61850.

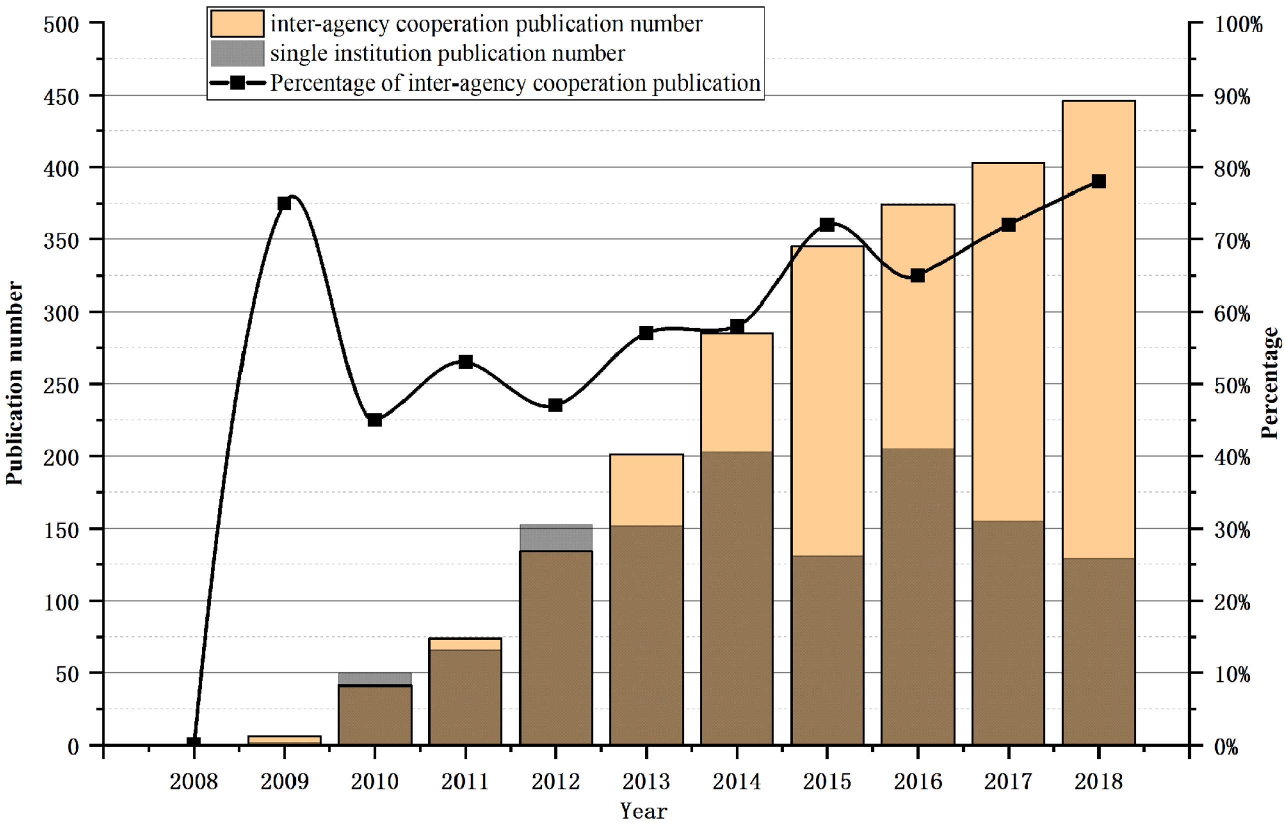

FIGURE 5 | Annual literature distribution of inter-agency and single institution.

Author-Keywords Co-occurrence Analysis With Three Stages

Based on author-keywords co-occurrence research, from 2008 to 2018, a total of 7,886 author-keywords were used. A total of 6,308 author-keywords appeared only once, and 123 author-keywords had a frequency >10.

In section Annual Publications and Growth Trend, we divided 2008–2018 into three stages. In this section, we will examine the co-occurrence of keywords in these three stages to find the research hotspots in each stage. **Figures 8–10** show the analysis map of keyword co-occurrence in the three stages, in which the node size is equal to the number of co-occurrences of keywords, and the larger the node, the more co-occurrences. The line between nodes represents the number of simultaneous occurrences of two keywords, and the thicker the line is, the more simultaneous occurrences of keywords were present.

Figure 8 is the keywords co-occurrence map of Chinese scholars' publications in the field of SG from 2008 to 2009, with keyword minimum occurrences as 1, 42 keywords were obtained. This diagram is composed of several independent parts, labeled as cluster 1–1 to cluster 1–7. Cluster 1–1 contains the keywords: machine learning, architecture, and human experience, which can be regarded as the application of artificial intelligence in the power system. Cluster 1–2, whose theme is grid planning and management, mainly contains keywords such as planning and management power grids. Cluster 1–3 focuses on renewable energy and sustainable development, including innovation, technology, and strategies. Cluster 1–4 focuses on integrating distributed energy to improve energy efficiency, mainly including keywords such as distributed generation and energy efficiency. Cluster 1–5 is the largest, and its theme is mainly based on artificial intelligence and the automatic fault identification of sensors, mainly including signal analysis, leak detection, signal detection, and other keywords. Cluster 1–6 contains only four keywords: smart substation, vision, smart control center, and smart transmission grid, which can be thought of as the clustering theme for intelligent control of power transmission and distribution. The last cluster 1–7 is mainly about the research of batteries. At this stage, China's SG has just started, and some attempts have been made at the technical level, but the correlation between different topics is weak.

SG in China is at an embryonic stage and is committed to solving the problem of renewable energy consumption. The research mainly focuses on the power grid itself, such as power grid dispatching, intelligent power transformation, and power grid planning. Meanwhile, battery, big data, and machine learning are proposed.

Figure 9 shows the keyword co-occurrence map of Chinese scholars' publications in the SG field from 2010 to 2014.

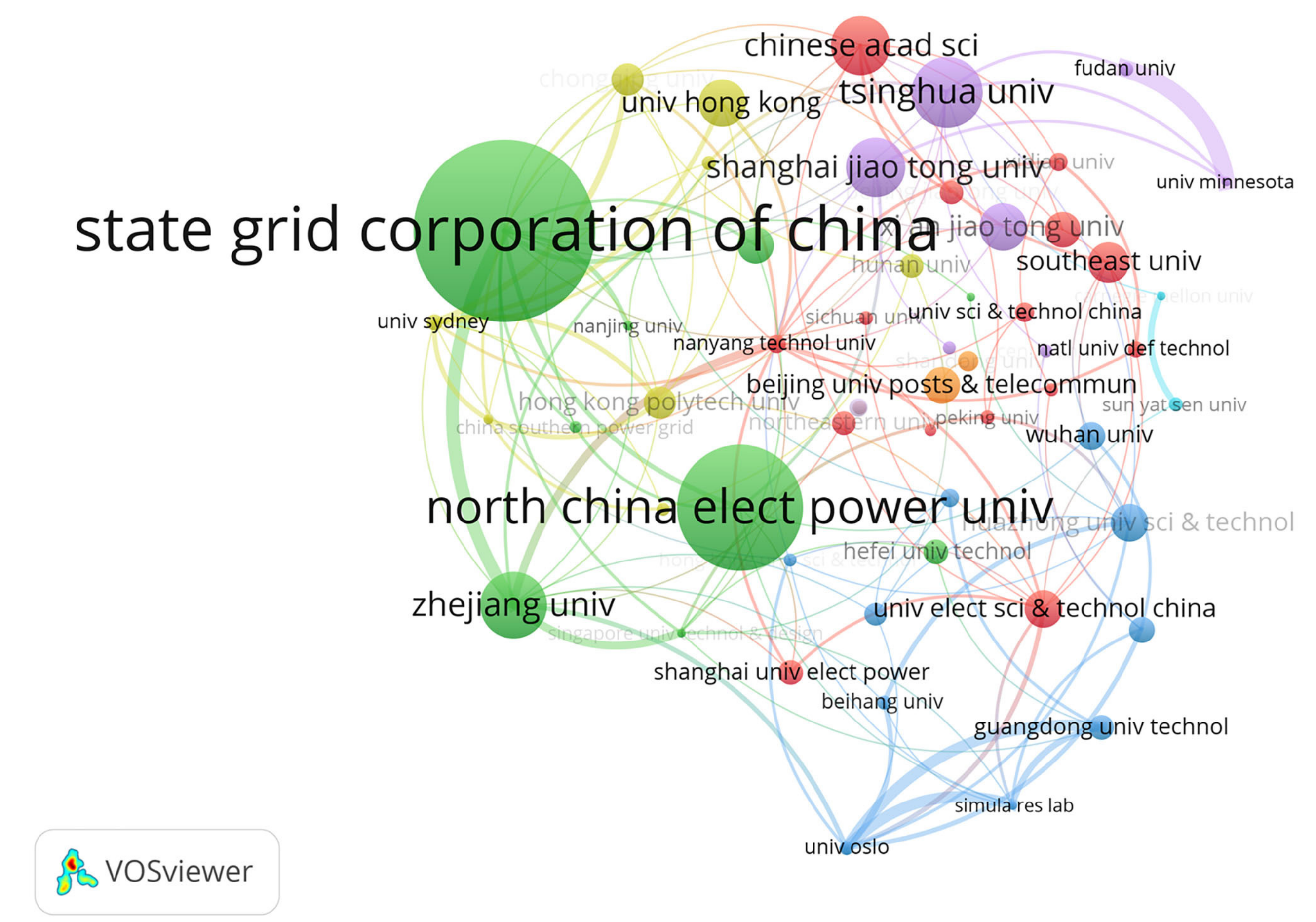

FIGURE 6 | Institution co-authorship network in China in the field of SG. VOSviewer parameters: counting method: full counting, minimum occurrences 10, a total of 61 institutions, normalization using LinLog/modularity method, layout parameters: attraction = 8, repulsion = 0.

Keywords at this stage are considered more closely related, indicating that the research in the SG field at this stage is more scattered and involves more fields. The top 10 keywords at this stage are demand response, EV, micro-grid, WSN, security, IoT, distribution generation, electricity, IEC61850, and distribution network. VOSviewer automatic clustering produces six clusters, labeled as cluster 2–1 to cluster 2–6. The largest cluster is cluster 2–1, which contains 29 keywords on the integration of distributed energy, including the following keywords: EV, micro-grid, distributed generation, and distribution network. The second-ranking cluster is cluster 2–2, which contains 28 keywords. The theme is DSM to improve energy efficiency. The main keywords include demand response, self-healing, and multi-agent. The next cluster is cluster 2–3, with 25 keywords. The theme is energy storage, batteries, real-time systems, scheduling, and delays. Cluster 2–4 is the technology base of SG, including IoT, WSN, sensors, reliability, quality of service, and other keywords. Security is becoming increasingly important in SG. Cluster 2–5 contains related topics and consists of 16 keywords, including intelligent electricity meter, cloud computing, and advanced metering infrastructure (AMI). The last cluster 2–6 focuses more on the establishment of SG standards, including smart grid standards such as IEC61850.

This stage belongs to the development period of SG in China. Driven by policies and technological progress, research keywords are exploding, which mainly reflect three research themes. Firstly, the SG technology foundation, such as IoT and WSN, are emerging in a concentrated manner, among which smart meters are the most represented in the IoT field. At the same time, studies on SG standards also began to appear, such as IEC61850 related research. Secondly, SG lead to system reconstruction of the power system, traditional power grid power generation, transmission, substation, power distribution, and utilization and scheduling boundaries become blurred, distributed energy, and micro-grid become relatively independent of the interconnected power system, EVs became the new way of energy storage, demand-side management combined with smart meters, have the potential to further enhance energy efficiency. SG has transformed the power system from a traditional independent link into an interconnected and mutually reinforcing internet.

Figure 10 shows the map of keyword co-occurrence of Chinese scholars in the SG field from 2015 to 2018. The minimum total number of occurrences was 5, and 139 keywords

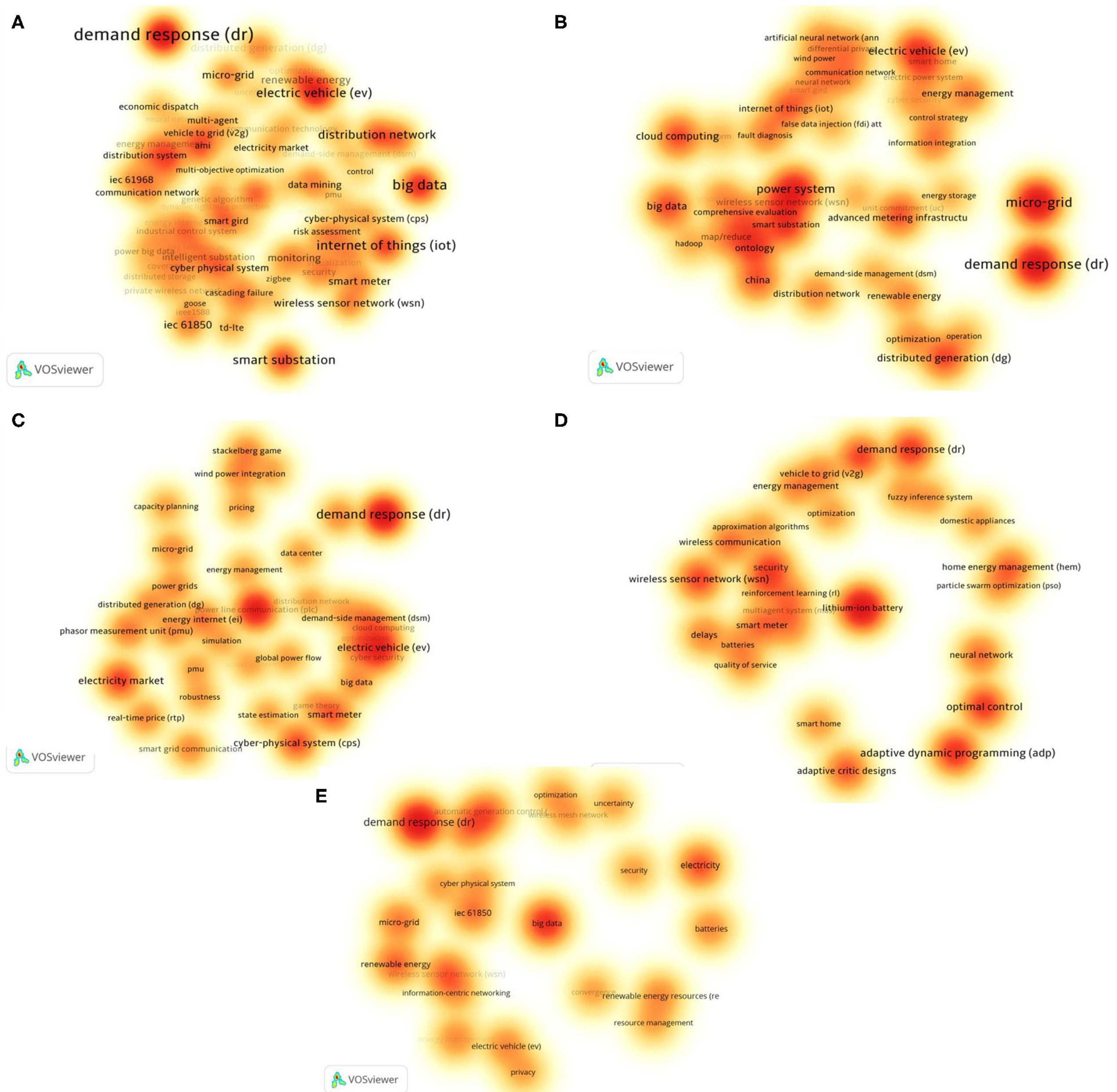

FIGURE 7 | Main research topics in the top five most productive institutions. Author-keywords are selected, counting method: full counting, search keywords such as smartgrid, smart grid, and smart grids are not included. The brighter the color, the more frequently the keyword appears. VOSviewer parameters: normalization method: LinLog/modularity, attraction = 8, repulsion = 1, clustering: resolution = 1, min. The cluster size: 1, density kernel width: item density = 1.18. **(A)** The research topic of the State Grid Corporation of China. A total of 1,802 keywords, minimum occurrence is 3, 131 keywords were obtained. **(B)** The research topic of North China Elect Power University. A total of 1,045 keywords, minimum occurrence is 3, 49 keywords were obtained. **(C)** The research topic of Tsinghua University. A total of 581 keywords, minimum occurrence is 3, 40 keywords were obtained. **(D)** The research topic of the Chinese Academy of Sciences. A total of 490 keywords, minimum occurrence is 3, 31 keywords were obtained. **(E)** The research topic of Shanghai Jiaotong University. A total of 613 keywords, minimum occurrence is 3, 26 keywords were obtained.

were obtained. In the third stage, the research mainly focuses on demand response, EV, big data, smart meter, micro-grid, WSN, security, IoT, optimization, cloud computing, and distributed generation. Similar to the previous stage, the keyword co-occurrence map in this stage is automatically clustered into 6 clusters, labeled as cluster 3–1 to cluster 3–6. The largest cluster is cluster 3–1, whose theme is similar to cluster 2–4. It is the technical basis of SG and includes 28 members, including WSN, IoT, smart meters, power line communication, and reliability. Cluster 3–2 covers all aspects of the power system, such as

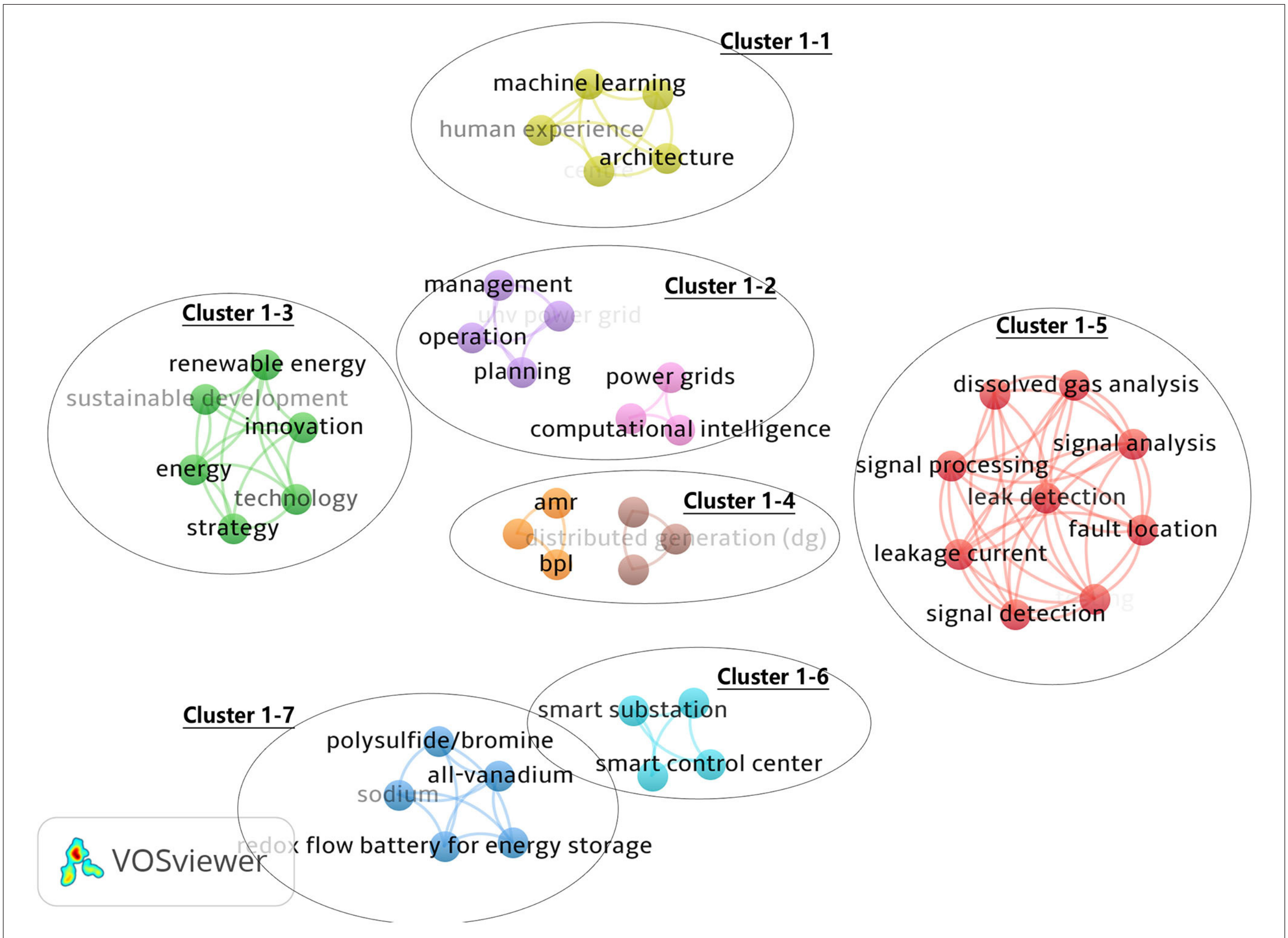

FIGURE 8 | Author-keyword co-occurrence network of SG publications in China, 2008–2009. A total of 42 keywords, VOSviewer parameters: minimum occurrence is 1, normalization method: LinLog/modularity, attraction = 3, repulsion = 0, clustering: resolution = 1 min. The cluster size: 1. Search keywords such as smartgrid, smart grid, and smart grids are not included.

demand response, energy storage, distributed network, power market, and micro-grid. Its theme is SG's impact on the power system. At this stage, EV becomes an independent cluster (cluster 3–3), indicating that EV are getting more attention because they are both an energy consumer and an energy provider, which is of great significance for energy management. Cluster 3–4 has the theme of power system optimization and contains 14 members. The main keywords are optimization, distributed control, economic dispatch, and load management. Cluster 3–5 is for big data processing; SG generates massive data, which need to be processed by relevant big data methods, including big data, load forecasting, and other keywords. Cluster 3–6 includes cyber-physical systems (CPS), smart substation, cyber-attack, cascading failure, state estimation, cyber security, and false data injection attack, the main theme for the physical information system and its safety.

The third stage is the stability period of China's SG development. Keywords in this stage are more detailed, as shown in **Figure 10**, the number of nodes with larger diameters have increased significantly, which can be roughly divided into three themes. The first is still the technical basis of SG. In this stage, in addition to the main keywords of the previous stage, such as IoT, WSN, and smart meter, some new features have emerged. If big data become a research hotspot, with the development of SG and more data collection, the problem of data analysis will be raised. At the same time, security has become another important topic. Sensor, data communication, system vulnerability, and other security problems are increasing in SG. Security and stability are the basic requirements of the power system. The second theme is the impact of SG on the power system. Keywords also appear in the trend of decentralization, such as renewable energy, energy management, and energy efficiency. China is facing the grim situation of greenhouse gas emissions. At the same time, the energy internet appears in this stage, which is the depth and development of SG. At this stage, research on EV and big data become independent clusters, and their research becomes more detailed.

Research in China's SG field has obvious interdisciplinary features. For example, the wireless transmission network is an important way to ensure data transmission. It belongs to the

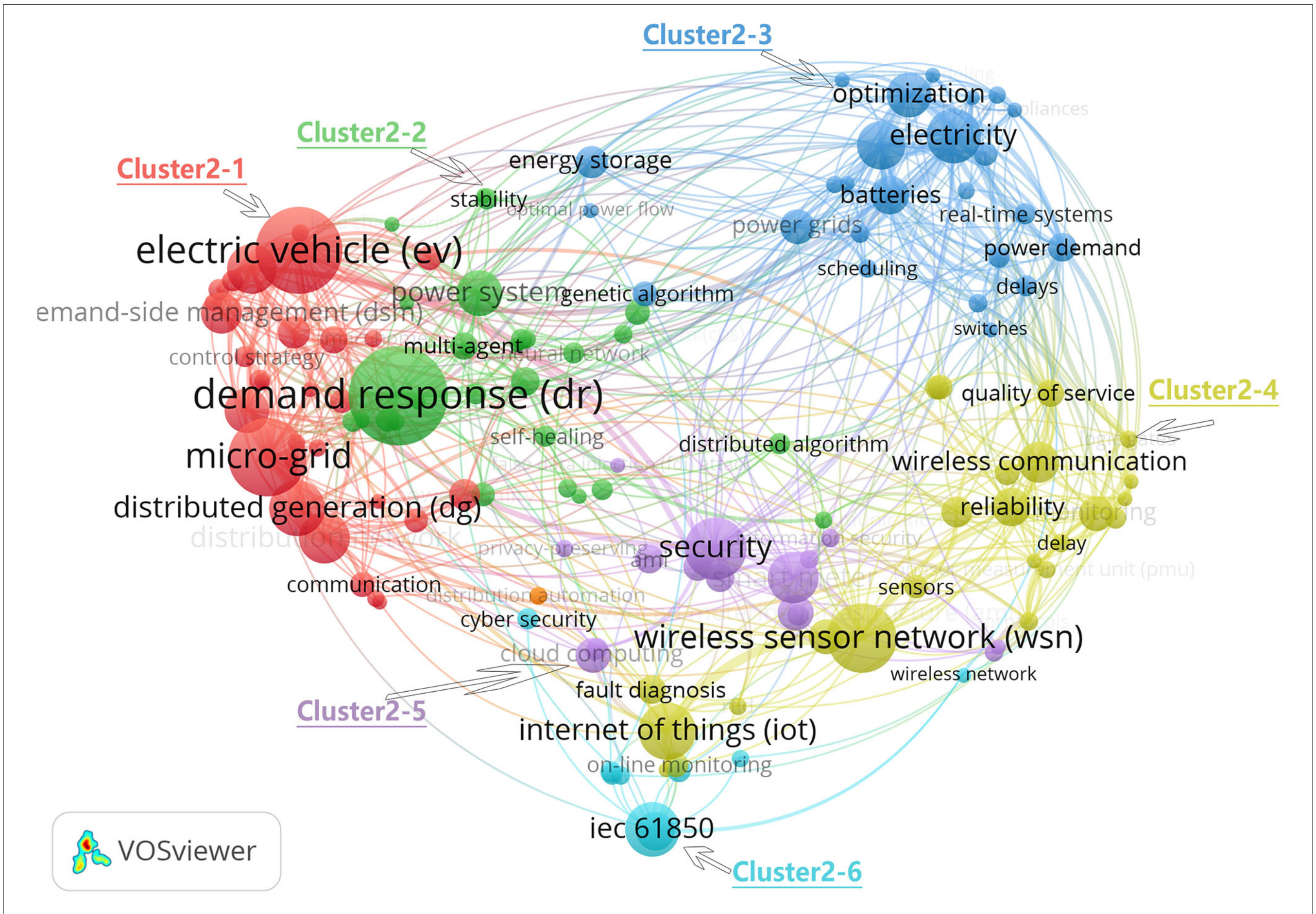

FIGURE 9 | Author-keywords co-occurrence network of SG publications in China, 2010–2014. A total of 3,185 keywords, VOSviewer parameters: minimum occurrence is 5, normalization method: LinLog/modularity, attraction = 1, repulsion = 0, clustering: resolution = 1, min. The cluster size: 5. Search keywords such as smartgrid, smart grid, and smart grids are not included.

application of data communication in the smart power grid. The IoT is an important carrier for smart devices to play their roles. It involves electronics, communication, and other fields. The power system is increasing the proportion of renewable energy, SG can partly solve the intermittent issues. Energy storage is an important method. How to plan as a whole and consider the power system under the influence of smart grid optimization is a problem that academic circles have been exploring, so in the future smart grid, the influence of traditional power systems will be bigger; China has the most complex power grid in the world, with complex power consumption and users. Demand response has always been a hot research topic. Chinese researchers have been seeking to realize optimal dispatching of the power grid through demand response, to guarantee the security and stability of the power grid.

The Alluvial Diagram of Author-Keywords

According to the clustering of keywords at different stages in **Figures 8–10**, the research hotspots at different stages can be roughly seen, but the development and evolution of research hotspots cannot be shown. Therefore, the alluvial diagram is used in this paper to show the evolution of hotspots at different stages.

Looking for changes in the scientific structure is important for understanding the development of science. Alluvial diagrams are such a tool that reveal stories in the network data and allow us to connect structural and functional changes (Rosvall and Bergstrom, 2010). This article utilizes MapEquation's alluvial diagrams to understand the changing trends of research hotspots in the SG field. Firstly, the data set was pre-processed in CiteSpace (Chen, 2013). It was divided into three stages, 2008–2009, 2010–2014, and 2015–2018, respectively. Three.net files were generated and used to import MapEquation's alluvial diagrams APP. Alluvial diagrams illustrated in **Figures 11A–C** were finally generated after computing clusters to simplify the network (Rosvall and Bergstrom, 2008)[1].

In **Figure 11**, every equivalently colored block in the alluvial diagram presents a cluster in the networks. The change of the cluster structure from one time period to the next period was

[1]Specific operational method and file format can refer to the website https://www.mapequation.org/apps/MapGenerator.html

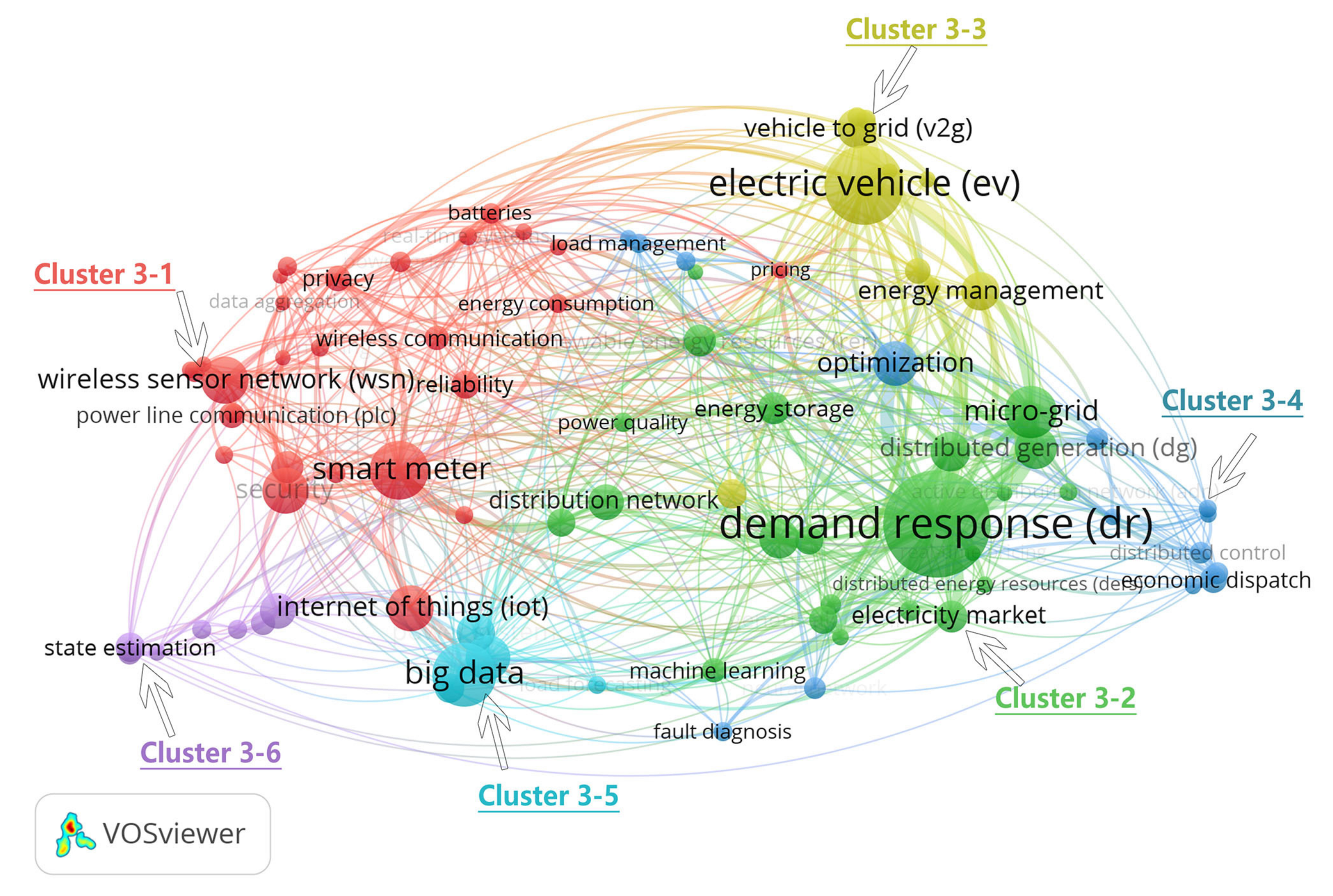

FIGURE 10 | Author-keyword co-occurrence network of SG publications in China, 2015–2018. A total of 5,457 keywords, VOSviewer parameters: minimum occurrence is 10, normalization method: LinLog/modularity, attraction = 1, repulsion = 0, clustering: resolution = 1 min. The cluster size: 1. Search keywords such as smartgrid, smart grid, and smart grids are not included.

represented by the merge and divergence in the strips connecting the time 1 and time 2 blocks (Rosvall and Bergstrom, 2010). For example, **Figure 11A** shows how SG in China evolved from a relatively independent topic to a cluster of many research topics between 2008 and 2018.

The height of the block reflects the size of the cluster, the clusters are ordered from bottom to top by size, with mutually non-significant clusters placed together and separated by a third of the standard spacing. Here we will prefix # with the cluster name to differentiate it with the subset. For example, in the 2015–2018 column in **Figure 11A**, the cluster #battery at the bottom is the largest, so it has the highest block. Cluster #SG is the second largest, ranking on top of the battery, with smaller blocks. The height of a stream field, going from the subset of a cluster in one column to the subset of a cluster in the adjacent column, which represents the total size of the nodes that make this particular transition. By following all stream fields from a cluster to an adjacent column, it is possible to study in detail the mergers with other clusters and the significant transitions. The block representing cluster #SG is composed of many streams, the largest of which is the orange stream at the bottom, which represents the highest proportion of PageRank for this node (Lambiotte and Rosvall, 2012). The orange stream flows from 2008 to 2009 through 2010–2014 and finally flows into 2015–2018, indicating that this research topic runs through three stages and has been a hot research topic. Similarly, **Figure 11B** shows the composition and flow of cluster #battery, while **Figure 11C** shows other clustering situations from 2015 to 2018, such as #big data, #automated metering infrastructure (AMI), #state estimation, #optimization, #distributed algorithm, and #communication.

Figure 11A shows the development process of cluster #SG. In the 2008–2009 stage, #SG only includes two subsets: SG and energy efficiency (EE). Among them, subset SG is transferred to the cluster in the stage 2010–2014 while EE is transferred to the other cluster #energy management. The subset of the new cluster #SG in the 2015–2018 phase has significantly increased. In addition to the orange SG stream, the subset also includes subsets who are transferred from the other cluster in the 2010–2014 phase. Subset demand response and electricity market transfers from cluster #energy management. Subset distribution network, distribution generation, renewable energy, and subset game theory is transferred from cluster #DG.

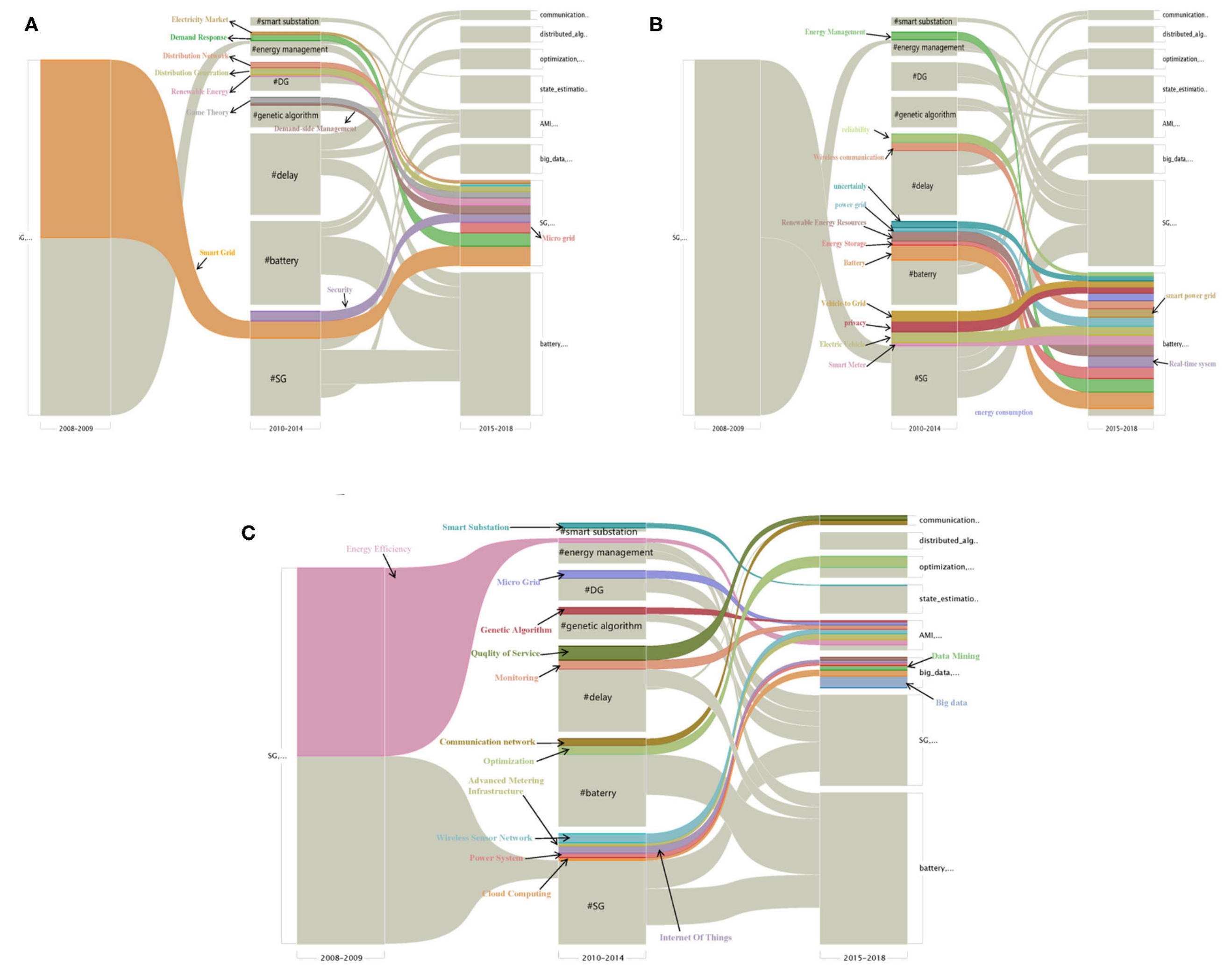

FIGURE 11 | The alluvial diagram of author-keywords in three stages. Significance clustering in 2008–2009, 2010–2014, and 2015–2018 takes up a column in the figure. Each block in a column represents a cluster, and the height of the block reflects the keywords flow for that cluster. Darker color was used to indicate the significant subset of each cluster. All keywords that are clustered in the subset of others in 2015–2018 are colored to highlight the fusion and formation of clusters. **(A)** Flow of keywords in SG clustering. **(B)** Flow of keywords in battery clustering. **(C)** Flow of keywords in other clustering.

Similarly, **Figure 11B** shows the data flow of cluster #battery in 2015–2018, which is mainly composed of streams imported from 2010–2014. For example, vehicle to grid, privacy, EV, and smart meter are imported from the subset in the cluster #SG. Power grid, renewable energy resources, energy storage, and battery are imported from the subset in cluster #SG. Besides, the subset of cluster #energy management and the subset of cluster #delay wireless communication also converge to the cluster #battery in the last stage.

Figure 11C shows the inflow of subsets of the last few clusters, such as #communication network, #state estimation, #optimization, #AMI, and #big data. #big data consists of a lot of subsets: big data, cloud computing, data mining, IoT, machine learning, and power system. Among them, only IoT, power system, and cloud computing are transferred from phase 2010–2014, while the rest of the topics all involve the new subset. #AMI mostly flows from the previous phase, genetic algorithm, micro-grid (#DG), monitoring, WSN, AMI, energy efficiency (#energy management), the subset of the newly emerged subset is load management. #communication network consists of the quality of service (#delay) and communication network transfer in the former phase and also includes one subset of the newly emerged subset PLC. #distributed algorithm is a new cluster that includes two subsets: distributed algorithm and economic dispatch. #optimization includes four subsets: optimization (#battery), multi-agent system, control, and neural network. #state estimation consists mainly of a new subset including the state estimation, false data injection attack, CPS, cyber-attack, system, and smart substation (#smart substation).

SG's keywords at the three stages of development are more refined, in the first stage there are only two clusters, new clusters appear in the second phase, such as #battery, #delay, #based algorithm, #distributed generation, and #smart substation, together with the existing similar findings in section Author-Keywords Co-occurrence Analysis With Three Stages, the stage of SG has a great influence on the traditional power system and begins to change each link in power system research. In the third stage, these clusters are re-decomposed and combined to generate the new clusters #big data, #AMI, and #optimization. In addition to paying attention to the impact of the smart power grid on the traditional power grid, they also turn to intelligent equipment, data processing, and decision optimization.

CONCLUSION

We applied different bibliometric methods and visualization tools to analyze the basic situation of China's SG, and tried to answer the questions raised in the first part by presenting bibliometric results.

Under the pressure of clean energy utilization and climate change, the development of SG has attracted the attention of various countries. This paper adopts a scientific visualization method to analyze 3,558 WOS literature records and uses scientometric techniques, such as co-authorship and co-occurrence, through which the status quo and trends of the SG field in China can be deeply understood.

(1) The basics of SG in China: The publication volume presents an overall rising trend, but fluctuates after 2015. The research literature of Chinese scholars in SG is expected to continue to grow in the coming years. At the same time, from the perspective of discipline distribution, the literature in the SG field has obvious interdisciplinary characteristics. We found that the top 10 institutions in the SG field published half of the literature in the data set.

2) Cooperation on SG research in China. China has cooperative relations with 52 countries in the SG field, and the cooperation ratio among countries shows an increasing trend. The proportion of collaborative papers continues to rise and so does the proportion of the formation of several major partnerships.

(3) Research hotspots of five productive institutions. SGCC, which has the largest number of published documents, is China's largest power grid service provider and is the main driver of SG development. It is committed to building an SG system based on power flow, information flow, and business flow of the entire power system. While increasing the consumption of renewable energy, SGCC plans to build a strong smart grid based on UHV, achieve coordinated development of power grids at all levels through information flow, and use it as a support to realize the integration of power generation, transmission, substation, distribution, power consumption, and dispatching business flow development. Its research topics are the most scattered and involve all aspects of SG research. In addition to SGCC, North China Electric Power University has a wide range of research hotspots, such as distributed power generation and micro-grid, demand response, EV, big data, and cloud computing. The research hotspots of Tsinghua University, CAS, and Shanghai Jiaotong University are relatively concentrated. In addition to demand response, Tsinghua University pays more attention to the cyber-physical system (CPS), power line communication (PLC), and EV. CAS pays more attention to power system planning and optimization, neural network, home energy management, and lithium-ion batteries. Shanghai Jiaotong University pays more attention to automatic generation control and big data.

(4) Research hotspots and development trends. SG research has interdisciplinary characteristics, which is the integration of power system, engineering technology, and information communication technology. The research topics include engineering technology development, power system reconstruction, and information communication development. In terms of engineering technology, the development of battery technology has driven the development of EV and energy storage. Hot keywords include the battery, EV, energy storage, and lithium-ion battery. At the same time, the development of sensor technology provides technical support for data collection. Related keywords include IoT, sensors, smart meter, and AMI. Power system reconfiguration driven by SG thus affects all aspects of the power system research. Related keywords include micro-grid, distributed generation, smart substation, v2g, energy internet, and optimization. The development of information and communication technology is the foundation of the development of SG, and its main research hotspots include WSN, wireless communication, security, and big data. The development of China's SG research has evolved from a relatively single theme to the application of information and communication technology to integrate power system-wide research topics, including distributed power generation, smart grid, smart power transformation, smart dispatch, and demand response.

In the next step, further research into SG in China needs to focus on the following tasks: hardware design based on the IoT, network construction based on wireless transmission network, optimal dispatching of the power system based on demand response, improvement of energy efficiency, and the construction of regional and national smart grids.

Strengths and Limitations

Quantitative analysis of a certain field or discipline based on bibliometric is a method that has gradually emerged in recent years. Although this method cannot be applied to accurately summarize the development of the discipline or field, it can find the basic characteristics and hotspots of the research of the discipline or field from a macro perspective. According to our search, this is the first bibliometric analysis of SG which uses alluvial diagrams to analyze the trend of keywords.

The research has provided valuable information for SG researchers, practitioners, and government institutions in China, meanwhile, the visualized map has provided valuable insight and an in-depth understanding of the key institutions, institutional and national cooperation, the current state of the research field, and the development trend.

The results of this study will be applied to (1) when government agencies and business organizations are formulating policies, consulting, and determining research cooperation; (2) when graduate students start to determine the current development situation of the gaps in understanding in the SG field; (3) when scholars make an attempt to understand the research hotspots in the SG field in China or need to discuss these topics with other scholars and seek potential cooperation.

There were some limitations in this study deserving our attention. First of all, our analysis is based on the analysis of the sample set, rather than the research results of the whole SG field in China. The SG field is developing rapidly and tends to be interdisciplinary. We are trying to extend it to the whole WOS library, however, the inclusion of the Emerging Sources Citation Index (ESCI) database starting in 2015, is different from the analysis period in this paper from 2008 to 2018, therefore this database was excluded in the analysis. In the subsequent research, the Scopus, Google Scholar, and other databases can be further taken into consideration for more comprehensive analysis in this field. Anyhow, the WOS database is the most comprehensive and widely used data at present, which is still meaningful for understanding the research status and development of the SG field in China. Secondly, this paper mainly considers SG research hotspots and its evolution in China from a macro perspective, but does not consider the relationship network of cooperation and co-citation between the authors, such as relationships regarding their common affiliations, academic supervisor-student relationships, and undoubtedly working experience is also important for understanding the development of SG research in China. Further research may consider the microscopic analysis of the authors and their cooperation.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/supplementary material.

AUTHOR CONTRIBUTIONS

CW, TL, and XD contributed to the conception and design of the study. XD organized the database. CW wrote the first draft of the manuscript. All authors contributed to manuscript revision and read and approved the submitted version.

FUNDING

This study was supported by the National Natural Science Foundation of China (72074212) and the Fundamental Research Funds for the Central Universities under grant agreement No. 2020ZDPYMS42.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frma.2020.551147/full#supplementary-material

REFERENCES

Baumeister, T. (2010). *Literature Review on Smart Grid Cyber Security. Simulation*. Available online at: http://csdl.ics.hawaii.edu/techreports/10-11/10-11.pdf (accessed May 10, 2020).

Callon, M., Courtial, J. P., Turner, W. A., and Bauin, S. (1983). From translations to problematic networks: an introduction to co-word analysis. *Soc. Sci. Inform.* 22, 191–235. doi: 10.1177/053901883022002003

Chen, C. (2013). "Trajectories of Search," in *Mapping Scientific Frontiers: The Quest for Knowledge Visualization*, ed. C. Chen (London: Springer), 143–161. doi: 10.1007/978-1-4471-5128-9_4

Chen, C. (2017). Science mapping: a systematic review of the literature. *J. Data Inform. Sci.* 2, 1–40. doi: 10.1515/jdis-2017-0006

Chen, M., Mao, S., and Liu, Y. (2014). Big data: a survey. *Mob. Netw. Appl.* 19, 171–209. doi: 10.1007/s11036-013-0489-0

Cheng, F., and Chen, J. (2012). Metal-air batteries: from oxygen reduction electrochemistry to cathode catalysts. *Chem. Soc. Rev.* 41, 2172–2192. doi: 10.1039/c1cs15228a

Deng, R., Yang, Z., Chow, M. Y., and Chen, J. (2015). A survey on demand response in smart grids: mathematical models and approaches. *IEEE Trans. Industr. Inform.* 11, 570–582. doi: 10.1109/TII.2015.2414719

Ding, Y., Rousseau, R., and Wolfram, D. (2014). *Measuring Scholarly Impact: Methods and Practice*. Springer International Publishing. p. 346.

El-Hawary, M. E. (2017). "The Smart grid-state-of-the-art and future trends," in *2016 Eighteenth International Middle East Power Systems Conference* (Cairo: MEPCON). doi: 10.1109/MEPCON.2016.7836856

Fadaeenejad, M., Saberian, A. M., Fadaee, M., Radzi, M. A. M., Hizam, H., and Abkadir, M. Z. A. (2014). The present and future of smart power grid in developing countries. *Renew. Sustain. Energ. Rev.* 29, 828–834. doi: 10.1016/j.rser.2013.08.072

Gao, J., Xiao, Y., Liu, J., Liang, W., and Chen, C. L. P. (2012). A survey of communication/networking in smart grids. *Future Gener. Comput. Syst.* 28, 391–404. doi: 10.1016/j.future.2011.04.014

Garfield, E. (1979). *Citation Indexing: Its Theory and Applications in Science, Technology, and Humanities. 1st Edn,* ed. E. Garfield (New York, NY: John Wiley & Sons, Ltd).

Garfield, E. (1983). Citation indexes for science a new dnenaion in documentation through associationof ideas. *Science* 1, 468–471.

Goran, S. (2010). SSCI and its impact factors: a "prisoner's dilemma"? *Eur. J. Mark.* 44, 23–33. doi: 10.1108/03090561011008583

Gungor, V. C., Sahin, D., Kocak, T., Ergut, S., Buccella, C., Cecati, C., et al. (2011). Smart grid technologies: communication technologies and standards. *IEEE Trans. Ind. Inform.* 7, 529–539. doi: 10.1109/tii.2011.2166794

Guo, Y. M., Huang, Z. L., Guo, J., Li, H., Guo, X. R., and Nkeli, M. J. (2019). Bibliometric analysis on smart cities research. *Sustainability* 11:3606. doi: 10.3390/su11133606

He, Q. (1999). Knowledge discovery through co-word analysis. *Libr. Trends* 48, 133–159.

Hirsch, J. (2005). An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. U.S.A.* 102, 16569–16572. doi: 10.1073/pnas.0507655102

Hossain, E., Khan, I., Un-Noor, F., Sikander, S. S., and Sunny, M. S. H. (2019). Application of big data and machine learning in smart grid, and associated security concerns: a review. *IEEE Access* 7, 13960–13988. doi: 10.1109/ACCESS.2019.2894819

Hossain, M. S., Madlool, N. A., Rahim, N. A., Selvaraj, J., Pandey, A. K., and Khan, A. F. (2016). Role of smart grid in renewable energy: an overview. *Renew. Sust. Energ. Rev.* 60, 1168–1184. doi: 10.1016/j.rser.2015.09.098

Hu, M., Pang, X., and Zhou, Z. (2013). Review recent progress in high-voltage lithium ion batteries. *J. Power Sources* 237, 229–242. doi: 10.1016/j.jpowsour.2013.03.024

Kabalci, Y. (2016). A survey on smart metering and smart grid communication. *Renew. Sust. Energ. Rev.* 57, 302–318. doi: 10.1016/j.rser.2015.12.114

Lambiotte, R., and Rosvall, M. (2012). Ranking and clustering of nodes in networks with smart teleportation. *Phys. Rev. E* 85:056107. doi: 10.1103/PhysRevE.85.056107

Li, F., Qiao, W., Sun, H., Wan, H., Wang, J., Xia, Y., et al. (2010). Smart transmission grid: Vision and framework. *IEEE Trans. Smart Grid* 1, 168–177. doi: 10.1109/TSG.2010.2053726

Lin, J., Yu, W., Zhang, N., Yang, X., Zhang, H., and Zhao, W. (2017). A survey on internet of things: architecture, enabling technologies, security and privacy, and applications. *IEEE Internet Things J.* 4, 1125–1142. doi: 10.1109/JIOT.2017.2683200

Liu, J., Xiao, Y., Li, S., Liang, W., and Chen, C. L. P. (2012). Cyber security and privacy issues in smart grids. *IEEE Commun. Surv. Tutor.* 14, 981–997. doi: 10.1109/SURV.2011.122111.00145

Liu, X., Bollen, J., Nelson, M. L., and Van De Sompel, H. (2005). Co-authorship networks in the digital library research community. *Inform. Process. Manag.* 41, 1462–1480. doi: 10.1016/j.ipm.2005.03.012

Mahmood, A., Javaid, N., and Razzaq, S. (2015). A review of wireless communications for smart grid. *Renew. Sust. Energ. Rev.* 41, 248–260. doi: 10.1016/j.rser.2014.08.036

Merigó, J. M., Mulet-Forteza, C., Valencia, C., and Lew, A. A. (2019). Twenty years of tourism Geographies: a bibliometric overview. *Tour. Geogr.* 21, 881–910. doi: 10.1080/14616688.2019.1666913

Montoya, F. G., Montoya, M. G., Gómez, J., Manzano-Agugliaro, F., and Alameda-Hernández, E. (2014). The research on energy in Spain: a scientometric approach. *Renew. Sust. Energ. Rev.* 29, 173–183. doi: 10.1016/j.rser.2013.08.094

Mulet-Forteza, C., Genovart-Balaguer, J., Merigó, J. M., and Mauleon-Mendez, E. (2019). Bibliometric structure of IJCHM in its 30 years. *Int. J. Contemp. Hosp. Manag.* 31, 4574–4604. doi: 10.1108/IJCHM-10-2018-0828

National Development and Reform Commission and National Energy Administration (2015). *Guiding Opinions on Promoting the Development of Smart Grid. National Development and Reform Commission*. Available online at: http://www.nea.gov.cn/2015-07/07/c_134388049.htm (accessed July 15, 2020).

National Development and Reform Commission and National Energy Administration (2016). *13th Five-Year Plan for Electric Power Development* (2016–2020). Available online at: http://zfxxgk.ndrc.gov.cn/web/iteminfo.jsp?id=398 (accessed July 20, 2020).

National Energy Administration (2010). *Notice on the establishment of the National Smart Grid Standardization Overall Working Group*. Available online at: http://www.nea.gov.cn/2011-10/12/c_131187554.htm (accessed July 20, 2020).

National Energy Administration and Ministry of Justice (2020). *Opinions on Speeding up the Establishment of Green Production and Consumption Regulations and Policy System*. Available online at: https://www.ndrc.gov.cn/xxgk/zcfb/tz/202003/t20200317_1223470.html (accessed July 19, 2020).

Olawumi, T. O., and Chan, D. W. M. (2018). A scientometric review of global research on sustainability and sustainable development. *J. Clean. Prod.* 183, 231–250. doi: 10.1016/j.jclepro.2018.02.162

Olawumi, T. O., Chan, D. W. M., and Wong, J. K. W. (2017). Evolution in the intellectual structure of BIM research: a bibliometric analysis. *J. Civil Eng. Manag.* 23, 1060–1081. doi: 10.3846/13923730.2017.1374301

Pan, H., Hu, Y. S., and Chen, L. (2013). Room-temperature stationary sodium-ion batteries for large-scale electric energy storage. *Energ. Environ. Sci.* 6, 2338–2360. doi: 10.1039/c3ee40847g

Perianes-Rodriguez, A., Waltman, L., and van Eck, N. J. (2016). Constructing bibliometric networks: a comparison between full and fractional counting. *J. Informetr.* 10, 1178–1195. doi: 10.1016/j.joi.2016.10.006

Ponce-Jara, M. A., Ruiz, E., Gil, R., Sancristóbal, E., Pérez-Molina, C., and Castro, M. (2017). Smart grid: assessment of the past and present in developed and developing countries. *Energy Strategy Rev.* 18, 38–52. doi: 10.1016/j.esr.2017.09.011

Rahimi-Eichi, H., Ojha, U., Baronti, F., and Chow, M. Y. (2013). Battery management system: an overview of its application in the smart grid and electric vehicles. *IEEE Ind. Electron. Mag.* 7, 4–16. doi: 10.1109/MIE.2013.2250351

Rahman, A. I. M. J., Guns, R., Rousseau, R., and Engels, T. C. E. (2017). Cognitive distances between evaluators and evaluees in research evaluation: a comparison between three informetric methods at the journal and subject category aggregation level. *Front. Res. Metr. Anal.* 2:6. doi: 10.3389/frma.2017.00006

Rosvall, M., and Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proc. Natl. Acad. Sci. U.S.A.* 105, 1118–1123. doi: 10.1073/pnas.0706851105

Rosvall, M., and Bergstrom, C. T. (2010). Mapping change in large networks. *PLoS ONE* 5:e8694. doi: 10.1371/journal.pone.0008694

Shi, Y., and Liu, X. (2019). Research on the literature of green building based on the web of science: a scientometric analysis in citespace (2002–2018). *Sustainability* 11:3716. doi: 10.3390/su11133716

Strasser, T., Andrén, F., Kathan, J., Cecati, C., Buccella, C., Siano, P., et al. (2015). A review of architectures and concepts for intelligence in future electric energy systems. *IEEE Trans. Industr. Electron.* 62, 2424–2438. doi: 10.1109/TIE.2014.2361486

Su, W., Eichi, H., Zeng, W., and Chow, M. Y. (2012). A survey on the electrification of transportation in a smart grid environment. *IEEE Trans. Industr. Inform.* 8, 1–10. doi: 10.1109/TII.2011.2172454

Subramanyam, K. (1983). Bibliometric studies of research collaboration: a review. *J. Inform. Sci.* 6, 33–38. doi: 10.1177/016555158300600105

Sun, Y., Zhao, L., Pan, H., Lu, X., Gu, L., Hu, Y. S., et al. (2013). Direct atomic-scale confirmation of three-phase storage mechanism in Li4Ti5 O12 anodes for room-temperature sodium-ion batteries. *Nat. Commun.* 4:1870. doi: 10.1038/ncomms2878

Tan, X., Li, Q., and Wang, H. (2013). Advances and trends of energy storage technology in Microgrid. *Int. J. Electr. Power Energy Syst.* 44, 179–191. doi: 10.1016/j.ijepes.2012.07.015

Tsui, K. M., and Chan, S. C. (2012). Demand response optimization for smart home scheduling under real-time pricing. *IEEE Trans. Smart Grid* 3, 1812–1821. doi: 10.1109/TSG.2012.2218835

Tu, C., He, X., Shuai, Z., and Jiang, F. (2017). Big data issues in smart grid – a review. *Renew. Sust. Energ. Rev.* 79, 1099–1107. doi: 10.1016/j.rser.2017.05.134

Tuballa, M. L., and Abundo, M. L. (2016). A review of the development of Smart Grid technologies. *Renew. Sust. Energ. Rev.* 59, 710–725. doi: 10.1016/j.rser.2016.01.011

Usman, A., and Shami, S. H. (2013). Evolution of communication technologies for smart grid applications. *Renew. Sust. Energ. Rev.* 19, 191–199. doi: 10.1016/j.rser.2012.11.002

van Eck, N. J., and Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* 84, 523–538. doi: 10.1007/s11192-009-0146-3

Varaiya, P. P., Wu, F. F., and Bialek, J. W. (2011). Smart operation of smart grid: risk-limiting dispatch: ways of managing energy systems without endangering reliability, while utilizing many intermittent resources, are discussed in this paper. *Proc. IEEE* 99, 40–57. doi: 10.1109/JPROC.2010.2080250

Wang, Q., Zhang, C., Ding, Y., Xydis, G., Wang, J., and Østergaard, J. (2015). Review of real-time electricity markets for integrating distributed energy resources and demand response. *Appl. Energy* 138, 695–706. doi: 10.1016/j.apenergy.2014.10.048

Wang, Y. (2010). Research framework of technical standard system of strong & smart grid. *Automa. Electr. Power Syst.* 34, 1–6 [in Chinese].

Wang, Y., Yu, X., Xu, S., Bai, J., Xiao, R., Hu, Y. S., et al. (2013). A zero-strain layered metal oxide as the negative electrode for long-life sodium-ion batteries. *Nat. Commun.* 4:2365. doi: 10.1038/ncomms3858

Wu, C., Mohsenian-Rad, H., and Huang, J. (2012). Vehicle-to-aggregator interaction game. *IEEE Trans. Smart Grid* 3, 434–442. doi: 10.1109/TSG.2011.2166414

Xiao, J., Gu, W., Wang, C., and Li, F. (2012). Distribution system security region: definition, model and security assessment. *IET Gener. Transm. Distrib.* 6, 1029–1035. doi: 10.1049/iet-gtd.2011.0767

Yu, D., and He, X. (2020). A bibliometric study for DEA applied to energy efficiency: trends and future challenges. *Appl. Energy* 268:115048. doi: 10.1016/j.apenergy.2020.115048

Yu, D., Xu, Z., and Wang, X. (2020). Bibliometric analysis of support vector machines research trend: a case study in China. *Int. J. Mach. Learn. Cybern.* 11, 715–728. doi: 10.1007/s13042-019-01028-y

Yu, W., Liang, F., He, X., Hatcher, W. G., Lu, C., Lin, J., et al. (2017). A survey on the edge computing for the internet of things. *IEEE Access* 6, 6900–6919. doi: 10.1109/ACCESS.2017.2778504

Yu, Y., Yang, J., and Chen, B. (2012). The smart grids in China-a review. *Energies* 5, 1321–1338. doi: 10.3390/en5051321

Yuan, J., and Hu, Z. (2011). Low carbon electricity development in China - an IRSP perspective based on super smart grid. *Renew. Sust. Energ. Rev.* 15, 2707–2713. doi: 10.1016/j.rser.2011.02.033

Yuan, J., Shen, J., Pan, L., Zhao, C., and Kang, J. (2014). Smart grids in China. *Renew. Sust. Energ. Rev.* 37, 896–906. doi: 10.1016/j.rser.2014.05.051

Zhang, D., Han, X., and Deng, C. (2018). Review on the research and practice of deep learning and reinforcement learning in smart grids. *CSEE J. Power Energy Syst.* 4, 362–370. doi: 10.17775/CSEEJPES.2018.00520

Zhang, Y., Yu, R., Nekovee, M., Liu, Y., Xie, S., and Gjessing, S. (2012). Cognitive machine-to-machine communications: visions and potentials for the smart grid. *IEEE Netw.* 26, 6–13. doi: 10.1109/MNET.2012.6201210

Zhao, B., Ran, R., Liu, M., and Shao, Z. (2015). A comprehensive review of Li4Ti5O12-based electrodes for lithium-ion batteries: the latest advancements and future perspectives. *Mater. Sci. Eng. R Rep.* 98, 1–71. doi: 10.1016/j.mser.2015.10.001

Zhong, H., Xie, L., and Xia, Q. (2013). Coupon incentive-based demand response: theory and case study. *IEEE Trans. Power Syst.* 28, 1266–1276. doi: 10.1109/TPWRS.2012.2218665

Zhong, Q. C., Nguyen, P. L., Ma, Z., and Sheng, W. (2014). Self-synchronized synchronverters: inverters without a dedicated synchronization unit. *IEEE Trans. Power Electron.* 29, 617–630. doi: 10.1109/TPEL.2013.2258684

Zhou, K., Fu, C., and Yang, S. (2016). Big data driven smart energy management: From big data to big insights. *Renew. Sust. Energ. Rev.* 56, 215–225. doi: 10.1016/j.rser.2015.11.050

Zhu, J., Song, L. J., Zhu, L., and Johnson, R. E. (2019). Visualizing the landscape and evolution of leadership research. *Leadersh. Q.* 30, 215–232. doi: 10.1016/j.leaqua.2018.06.003

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Research Metrics and Analytics

ORIGINAL RESEARCH
published: 22 January 2021
doi: 10.3389/frma.2020.593494

Comparative Analysis of the Bibliographic Data Sources Dimensions and Scopus: An Approach at the Country and Institutional Levels

Vicente P. Guerrero-Bote[1], Zaida Chinchilla-Rodríguez[2], Abraham Mendoza[3] and Félix de Moya-Anegón[4]*

[1]*Departamento de Información y Comunicación, Universidad de Extremadura, Badajoz, Spain,* [2]*Instituto de Políticas y Bienes Públicos (IPP), Consejo Superior de Investigaciones Científicas (CSIC), Madrid, Spain,* [3]*Facultad de Ingeniería, Universidad Panamericana, Zapopan, Mexico,* [4]*SCImago Research Group, Granada, Spain*

This paper presents a large-scale document-level comparison of two major bibliographic data sources: Scopus and Dimensions. The focus is on the differences in their coverage of documents at two levels of aggregation: by country and by institution. The main goal is to analyze whether Dimensions offers as good new opportunities for bibliometric analysis at the country and institutional levels as it does at the global level. Differences in the completeness and accuracy of citation links are also studied. The results allow a profile of Dimensions to be drawn in terms of its coverage by country and institution. Dimensions' coverage is more than 25% greater than Scopus which is consistent with previous studies. However, the main finding of this study is the lack of affiliation data in a large fraction of Dimensions documents. We found that close to half of all documents in Dimensions are not associated with any country of affiliation while the proportion of documents without this data in Scopus is much lower. This situation mainly affects the possibilities that Dimensions can offer as instruments for carrying out bibliometric analyses at the country and institutional level. Both of these aspects are highly pragmatic considerations for information retrieval and the design of policies for the use of scientific databases in research evaluation.

Keywords: Dimensions, Scopus, bibliographic data sources, database coverage, research evaluation, scientometrics, bibliometrics

OPEN ACCESS

Edited by:
Rafael Repiso, Universidad Internacional De La Rioja, Spain

Reviewed by:
Alberto Martín-Martín, University of Granada, Spain
M. J. Cobo, University of Cádiz, Spain

****Correspondence:***
Zaida Chinchilla-Rodríguez
zaida.chinchilla@csic.es

Specialty section:
This article was submitted to Research Assessment, a section of the journal Frontiers in Research Metrics and Analytics

Received: *10 August 2020*
Accepted: *10 November 2020*
Published: *22 January 2021*

Citation:
Guerrero-Bote VP, Chinchilla-Rodríguez Z, Mendoza A and de Moya-Anegón F (2021) Comparative Analysis of the Bibliographic Data Sources Dimensions and Scopus: An Approach at the Country and Institutional Levels. Front. Res. Metr. Anal. 5:593494. doi: 10.3389/frma.2020.593494

INTRODUCTION

As new multidisciplinary scientific bibliographic data sources are coming onto the market, there is growing interest in comparative studies looking at aspects of the coverage they offer. Scholarly databases have begun to play an increasingly important role in the academic ecosystem. There are several reasons for this, including burgeoning competitiveness in research, greater availability of data, and the need to justify the use of public funds. This context has driven the diversification of evaluations of publication and citation data use cases as well as of research use cases that have not been met by existing scholarly databases (Hook et al., 2018). Since bibliometric methods are used in multiple areas for a variety of purposes, especially research evaluation, the results they provide may vary depending on the representativeness of the database used (Mongeon and Paul-Haus, 2016; Huang et al., 2020). The new data sources can offer several benefits for research evaluators because they may have better coverage or have capabilities that make them a better fit for a given impact evaluation task, and they can reduce the cost of evaluations and make informal self-evaluations of impact possible for researchers who would not pay to access that kind of data (Thelwall, 2018). Given the potential value of these data sources for research evaluation, it is important to assess their key properties to better understand their strengths and weaknesses, in particular, to decide whether their data is sufficient in volume, completeness, and accuracy to be useful for scientists, policymakers, and other stakeholders.

Traditionally, the only homogeneous record of published research available when funders and governments sought additional information to help them make evidence-driven decisions was the Web of Science (WoS). The appearance of the Scopus database (Baas et al., 2020) and Google Scholar in 2004 as "competitors" to WoS, providing metadata on scientific documents and on citation links between these documents, led to an immense quantity of studies focused on comparative analyses of these other new bibliographic sources, the basic intention being to look for novel bibliometric opportunities that these tools might bring to the academic community and policymakers.

At that time, it appeared that Scopus and WoS had entered into head-on competition (Pickering, 2004), and any comparison of them called for the utmost care and methodological consistency. One large-scale comparison at the journal level was done using Ulrich's directory as the gold standard by Moya-Anegón et al. (2007). The results outlined a profile of Scopus in terms of its coverage by areas—geographic and thematic—and the significance of peer-review in the publications. Both of these aspects are of highly pragmatic significance for policymakers and the users of scientific databases. Years later, Mongeon and Paul-Haus (2016) revisited the issue and compared the coverage of WoS and Scopus to examine whether preexisting biases (such as language, geography, and theme) were still to be found in Scopus. They concluded that some biases still remained in both databases and stated that this should be taken into account in assessing scientific activities. For example, most languages and countries are underrepresented, which contributes to the known lack of visibility of research done in some countries. Hence, when using bibliometric methods for research evaluation, it is important to understand what each tool has to offer and what its limitations are and to choose the right tool for the task at hand before drawing conclusions for research evaluation purposes (Mongeon and Paul-Haus, 2016).

Google Scholar appeared to be an alternative to WoS and Scopus, but its suitability for research evaluation and other bibliometric analyses was called strongly into question. For a comprehensive review of this data source in research evaluation, we would refer to Martín-Martín et al. (2018a) and Martín-Martín et al. (2020).

At the beginning of 2018, Digital Science launched Dimensions, a new integrated database covering the entire research process from funding to research, from publishing results through attention, both scholarly and beyond, to commercial applications and policymaking, consistently matched in multiple dimensions (Adams et al., 2018). This new scholarly data source was created to overcome significant constraints of the existing databases. It sought to understand the research landscape through the lens of publication and citation data and help the academic community to formulate and develop its own metrics that can tell the best stories and give the best context to a line of research (Bode et al., 2019).

Previous studies have compared data quality between Dimensions and other data sources in order to evaluate its reliability and validity (Bornmann, 2018; Martín-Martín et al., 2018; Thelwall, 2018; Visser et al., 2020). Most of them have focused on publication and citation in specific thematic fields, but few of them have taken a global perspective. The findings of these studies in the field of Food Science show Dimensions to be a competitor to WoS and Scopus in making nonevaluative citation analyses and in supporting some types of formal research evaluations (Thelwall, 2018). Similarly, Martín-Martín et al. (2018b) conclude that Dimensions is a clear alternative for carrying out citation studies, being capable of rivalling Scopus. But the reliability and validity of its field classification scheme were questioned. This scheme is not based on journal classification systems as it is in WoS or Scopus, but on machine learning. This feature makes it desirable to undertake large-scale investigations in future studies to ensure that metrics such as the field-normalized citation scores presented in Dimensions and calculated based on its field classification scheme are indeed reliable (Bornmann, 2018).

A large-scale comparison of five multidisciplinary bibliographic data sources, including Dimensions and Scopus, was carried out recently by Visser et al. (2020). They used Scopus as the baseline for comparing and analyzing not just the different coverage of documents over time by document type and discipline but also the completeness and accuracy of the citation links. The results of this comparison shed light on the different types of documents covered by Dimensions but not by Scopus. These are basically meeting abstracts and other short items that do not seem to make a very substantial contribution to science. The authors concluded that differences between data sources should be assessed in accordance with the purpose for which the data sources are used. For example, it may be desirable to work within a more restricted universe of documents, such as a specific thematic field or a specific level of aggregation. This is the case with the study of Huang et al. (2020) which compared WoS, Scopus, and Microsoft Academic and their implications for the robustness of university rankings.

The present communication extends previous comparisons of Scopus by expanding the study set to include distinct levels of aggregation (by country and by institution) across a larger selection of characteristics and measures. A particular aim is to inquire closely into just how balanced Dimensions' coverage is compared with that of the Scopus database.

Objectives/Research Questions

The goal of this study was to compare Dimensions' coverage with that of Scopus at the geographic and institutional levels. The following research questions were posed:

(1) How comprehensive is Dimensions' coverage compared with that of Scopus in terms of documents?

(2) Are the distributions of publications by country and by institution in Dimensions comparable with those in Scopus?
(3) Are Dimensions' citation counts by country and by institution interchangeable with those of Scopus in the sense of their being strongly correlated?
(4) Is Dimensions a reliable new bibliometric data source at the country and institutional levels?

MATERIAL AND METHODS

Scopus is a scientific bibliography database created by Elsevier in 2004 (Hane, 2004; Pickering, 2004) which has been extensively characterized (Moya-Anegón et al., 2007; Archambault et al., 2009; Leydesdorff et al., 2010) and used in scientometric studies (Gorraiz et al., 2011; Jacso, 2011; Guerrero-Bote and Moya-Anegón, 2015; Moya-Anegón et al., 2018). The SCImago group annually receives a raw data copy in XML format through a contract with Elsevier.

In 2018, Digital Science published the Dimensions database with scientific publications and citations, grants, patents, and clinical trials (Hook et al., 2018; Herzog et al., 2020). Since then, there has been characterization published of it (Bornmann, 2018; Harzing, 2019; Visser et al., 2020). In the present study, we shall only consider the scientific publications.

Bibliographic databases often give bibliometric studies problems with author affiliations which usually do not include standardized names of institutions. One of the improvements that Dimensions incorporates is the mapping of author affiliations in documents to an entity list for organizations involved in research. This is the GRID (Global Research Identifier Database) system (Hook et al., 2018). This mapping is not an addition to but a replacement for author affiliations. If this mapping is rigorous and complete, it is an important improvement. But if the list of organizations or the mapping is incomplete, this could be a major problem because there would be loose documents without any possibility of associating them with institutions or countries, thus leaving the output of the institutions and countries affected incomplete.

The SCImago group has had the possibility of downloading a copy of Dimensions in Json format through an agreement with Dimensions Science.

From the Scopus and Dimensions data of April 2020, the SCImago group created a relational database for internal use that allows for massive computation operations that would otherwise be unfeasible.

Matching

For the analysis that was an objective of this study, it was necessary to implement a matching procedure between the Dimensions and Scopus databases. To this end, we applied the method developed in the SCImago group to match PATSTAT NPL references with Scopus documents (Guerrero-Bote et al., 2019). This method has two phases: a broad generation of candidate pairs, followed by a second phase of pair validation.

In this case, a modification was made, similar to that in Visser et al. (2020), in which not all the candidate pairs were generated at the same time. Instead, once there was a set of candidate pairs, a validation procedure was applied, accepting as valid the matches that exceeded a certain threshold. This reduced the combinatorial variability of the following generations of candidates. The pairs that did not exceed the threshold were not discarded but were saved in case at the end they were unpaired and were those with the greatest similarity.

In more detail, our procedure began with the normalization of the fields to facilitate pairing, although, unlike Visser et al. (2020), we did not stay exclusively with the numerical values of the volume, issue, or pages because at times those fields do not contain numerical values. This is the case with journals such as PLOS One or Frontiers, for instance.

Then we started to generate candidate pairs in phases. The phases were centered on the following conditions:

(1) One of these conditions:
   (1) Same year of publication, title with a high degree of similarity, and the same DOI.
   (2) Same year of publication, title with a high degree of similarity, and the same authors.
   (3) Same year of publication, title, and first author.

(2) One of these conditions:
   (1) Same year of publication and DOI.
   (2) Same year of publication, source (journal, proceeding, etc.), volume, and pages.
   (3) Same year of publication and coincidence in the first or last 20 characters of the title.
   (4) Same year of publication and authors.
   (5) Same year of publication and source.

As can be seen, there are conditions that include some previous phases. However, it should be borne in mind that each candidate pair generation phase is followed by a validation phase. So the first phases are quite specific; they generate a relatively small number of candidate pairs, most of which are accepted and come to constitute the majority of the definitively matched pairs. In this way, the lists of documents waiting to be matched are reduced, allowing for broader searches in the following phases without greatly increasing the computational cost. Logically, the percentage of success in the candidate pairs decreases from phase to phase.

For validation, all the reference's data were compared: DOI, year of publication, authors, title, publication, volume, issue, and pages. The last three were compared both numerically and alphanumerically. The comparison of each field generated a numerical score corresponding to the number of matching characters with some adjustments, for which the Levenshtein[1] distance was used

[1]In our case, we subtract the Levenshtein distance (multiplied by 1.3) from the number of characters in the largest of the fields to be compared, thus obtaining a number indicative of the number of matching characters between the fields (with a 30% penalty). Recall that the Levenshtein distance is the minimum number of single-character edits (insertions, deletions, or substitutions) required to change one string into the other.

as in Guerrero-Bote et al. (2019) and Visser et al. (2020). Once the coincidence score had been calculated in each field, we took the product to get the total score. The individual scores by field never have a zero value because that would mean the total score would be zero. In case of noncoincidence, the field score may be unity if the field is considered to be nonessential, 0.75 if it is considered to be important, etc. In either of the databases, the fields of some records may be empty. With this process, coincidence in several fields increases the total score geometrically rather than arithmetically.

Once the candidate pairs of a phase have been validated, we take as matched the pairs that obtain a total score greater than 1,000, and in which neither the Scopus nor the Dimensions record scores higher with any other pair. The total score threshold of 1,000 was set after sampling and verifying that under these conditions no mismatched pair was found.

Once the 5 phases had been carried out, a repechage operation was initiated for the rejected candidate pairs. This accepted pairs in which both components obtained a lower score in the rest of the pairs, down to a total score of 50. Also accepted were those in which the score was greater than 300, but one of the components had another pair with exactly the same score. This latter was done because both databases contain some duplicated records.

RESULTS

The Results of Matching

The general results are given in **Table 1**. It is true that, even though our study includes more years than that of Visser et al. (2020), it gives fewer matched documents for the period 2008–2017.

The number of matched pairs grows from year to year, and in Scopus, the percentage of matches also grows. This is not the case for Dimensions, however, due to the great growth this database experienced from year to year.

In summary, Dimensions' coverage is more than 25% greater than Scopus's, although there is a significant overlap in coverage between the two data sources. Almost three-quarters of the Scopus documents and more than half of the Dimensions documents match. The question now is to see if these percentage differences are maintained at levels of grouping of lower rank (countries and institutions).

The percentage of matching in Scopus by document type is presented in **Table 2**. The greatest percentages are in articles, reviews, letters, conference proceedings, errata, editorials, book chapters, short surveys, etc. (We have not listed some document types due to their low output.) For the primary output (articles, reviews, conference proceedings, and short surveys), the matching is over 75%.

Table 3 presents the same information, but for Dimensions. Articles and conference proceedings are the most matched types.

Figure 1 shows that the total and matched output distributed by country is systematically greater in Scopus than in Dimensions. The solid line represents the ideal positions of the countries if they had the same output in Scopus and Dimensions. It is noticeable at a glance that most countries appear above the solid line in the graph, indicating that the Scopus output by country tends to be greater than the Dimensions output.

Figure 2 shows the relationship of the output by institution between Dimensions and Scopus. The solid line represents the positions of the institutions if they had the same output in both databases. It is again noticeable at a glance that most institutions are above the solid line, indicating that there are more institutions with more output in Scopus than in Dimensions.

Figure 3 allows one to analyze the evolution of the average number of countries whose institutions correspond to the author's affiliations in the documents present in one or the other database. What most stands out in this graph is the difference between the two databases. The two sets of evolution should be very similar, and yet they are not. These differences remain stable over time and need to be confirmed with the data representing the evolution of the number of institutions that appear in the author's affiliations.

Figure 4 confirms, from the institutional perspective, the evolution of the average of institutions per document in the two databases and in the matched documents. The two sets of evolution reveal the average of institutional affiliations associated with the items in the four subsets of the two data sources. As can be seen, the comparison between the two graphical representations is consistent.

In order to check the influence of documents without a country on the averages presented in **Figures 3**, **4**, **Figure 5** shows the evolution of the percentage of items in the four subsets of documents that do not record any country for some reason. As can be seen in the figure, these percentages have a downwards trend over the years in the different subsets of documents, and the order of the curves is contrary to that in **Figures 3**, **4**, which is consistent from the perspective of data interpretation.

In general terms, one can say that the information about institutional affiliations that allows documents to be discriminated by country and institution has greater completeness in Scopus than in Dimensions. The case is similar when analyzing this same situation from the perspective of the matched documents. In terms of temporal evolution, despite the positive trend in the number of countries and institutions associated with the items in both databases, the difference between the two sources in this regard tends to be maintained over time.

A more detailed characterization of the Dimensions documents where no country affiliation data is available is provided in **Table 4**. The distribution of document types shows that there are distinct document types affected by this situation.

Using as a basis the citation data (**Figure 6**), it is easy to see that, both for total documents and for matched documents, the volume of citations in Scopus is in all cases greater than that of Dimensions, as noted previously by Visser et al. (2020). The case is similar when the problem is analyzed from the point of view of the citing date (**Figure 7**).

TABLE 1 | Overall results of the linking procedure.

Year	Total matches	% change	% matches Scopus	% matches Dimensions	Total Scopus	% change	Total Dimensions	% change
2003	1,102,377	—	70.01	56.05	1,571,723	—	1,966,869	—
2004	1,175,774	6.66	69.61	54.49	1,686,413	7.30	2,157,735	9.70
2005	1,295,013	10.14	67.34	57.17	1,920,131	13.86	2,265,278	4.98
2006	1,406,239	8.59	69.59	56.87	2,019,216	5.16	2,472,883	9.16
2007	1,485,168	5.61	69.95	53.45	2,124,118	5.20	2,778,498	12.36
2008	1,566,745	5.49	70.37	56.74	2,227,050	4.85	2,761,246	−0.62
2009	1,665,294	6.29	71.17	56.73	2,342,897	5.20	2,935,302	6.30
2010	1,768,496	6.20	71.78	57.65	2,465,117	5.22	3,067,425	4.50
2011	1,902,640	7.59	72.52	54.52	2,625,462	6.50	3,489,937	13.77
2012	1,986,358	4.40	72.13	55.19	2,755,115	4.94	3,599,181	3.13
2013	2,085,792	5.01	72.62	54.05	2,874,153	4.32	3,859,025	7.22
2014	2,147,442	2.96	73.6	52.77	2,922,477	1.68	4,069,795	5.46
2015	2,182,437	1.63	75.52	52.04	2,891,116	−1.07	4,193,437	3.04
2016	2,259,015	3.51	75.54	51.62	2,990,795	3.45	4,376,598	4.37
2017	2,357,244	4.35	75.22	49.94	3,133,127	4.76	4,720,253	7.85
2018	2,533,236	7.47	79.15	50.33	3,190,038	1.82	5,033,439	6.63
2019	2,659,664	4.99	81.03	51.60	3,270,544	2.52	5,154,828	2.41
Total	31,578,934	—	73.39	53.61	43,009,492	—	5,8,901,729	—

TABLE 2 | Scopus matching percentages by most frequent document type.

	ar	re	cp	sh	bk	ch	ed	le	no	er
2003	76.5	60.9	68.8	41.8	25.8	37.1	46.5	66.1	30.3	52.1
2004	76.6	63.4	65.3	28.9	2 6	38.3	48.7	73.6	31.4	56.9
2005	76.2	61.1	61.9	29.3	2 6	48.5	44.1	67.3	25.2	59.4
2006	76.5	66.8	63.9	35.7	21.6	49.5	43.8	62.9	38.1	6 2
2007	76.4	71.2	65.5	42.6	23.8	43.9	42.5	63.9	35.1	61.7
2008	77.4	76.6	63.9	40.2	20.1	40.6	43.9	64.6	39.6	60.9
2009	77.5	76.1	66.6	37.8	22.4	43.5	43.4	65.9	4 3	63.1
2010	77.9	76.3	69.3	35.1	25.2	42.5	44.3	67.7	43.9	65.6
2011	78.8	76.9	70.2	33.6	22.7	37.9	43.7	69.3	44.2	70.4
2012	79.4	80.1	6 8	37.9	22.3	36.4	43.4	72.3	44.1	7 0
2013	79.7	81.6	68.6	3 6	27.3	36.6	42.7	73.3	43.6	69.6
2014	79.7	80.3	67.2	41.9	25.8	44.7	45.8	74.6	46.3	68.3
2015	81.3	80.5	68.2	45.7	25.5	50.4	48.7	76.2	48.1	71.3
2016	8 2	79.9	67.3	43.7	26.8	4 2	49.1	77.9	51.3	73.8
2017	80.6	80.1	70.8	44.6	3 1	45.7	49.2	79.8	52.7	64.9
2018	83.6	83.7	73.4	47.8	29.5	57.5	52.8	83.5	59.6	71.4
2019	8 5	86.5	67.5	5 9	39.8	71.8	61.5	86.7	66.4	73.7
Total	79.7	75.4	67.7	38.6	25.7	44.7	46.8	72.3	4 5	6 7

AR, articles; RE, reviews; CP, conference proceedings; SH, short survey; BK, book chapter; ED, editorial; LE, letters; NO, note; ER, erratum.

TABLE 3 | Dimensions matching percentages by document type.

Year	Article	Book	Chapter	Monograph	Preprint	Proceeding
2003	61.04	0.22	18.75	8.98	3.87	72.23
2004	60.14	0.12	16.27	8.78	3.79	74.38
2005	61.16	0.26	26.04	14.68	4.29	78.99
2006	61.26	0.72	25.54	12.95	4.34	76.49
2007	61.64	0.82	15.81	14.58	4.67	80.36
2008	61.70	1.07	21.79	13.76	4.35	82.12
2009	61.69	1.03	24.28	14.31	4.71	79.56
2010	61.47	0.99	26.25	14.72	4.90	81.47
2011	61.16	0.93	18.52	12.06	4.64	80.82
2012	60.08	1.69	25.54	15.38	5.02	79.44
2013	59.28	2.12	24.41	16.84	5.64	78.32
2014	58.07	3.30	23.46	13.68	5.88	70.12
2015	56.83	3.61	28.99	12.52	6.04	58.01
2016	56.52	1.43	26.06	11.29	6.50	56.08
2017	54.84	1.22	25.16	14.58	6.37	55.87
2018	55.79	0.95	23.72	9.78	7.45	57.29
2019	57.59	0.15	20.92	5.23	8.00	57.11
Total	58.86	1.32	23.06	12.58	5.83	69.66

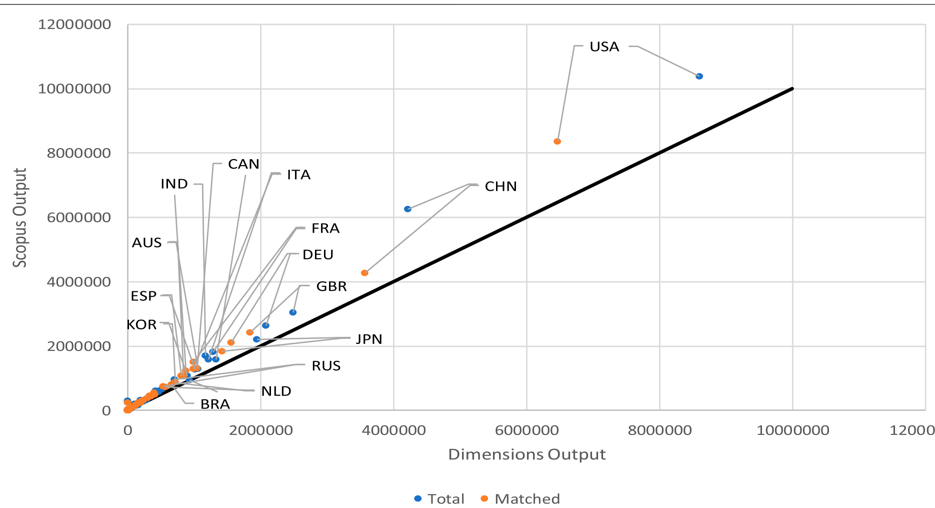

FIGURE 1 | Scatter plot of the total and matched Dimensions/Scopus output by country.

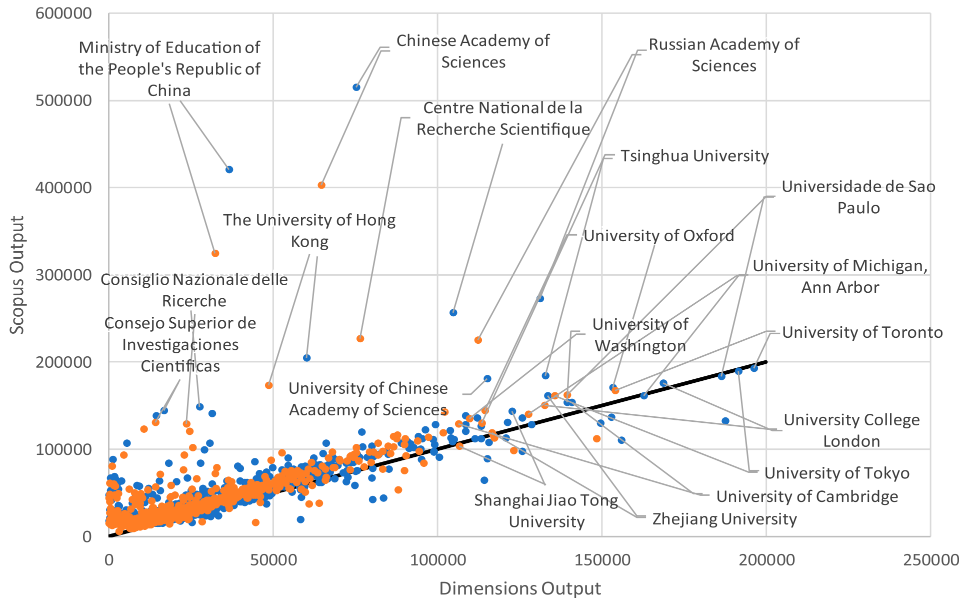

FIGURE 2 | Scatter plot of the total and matched Dimensions/Scopus output by institution.

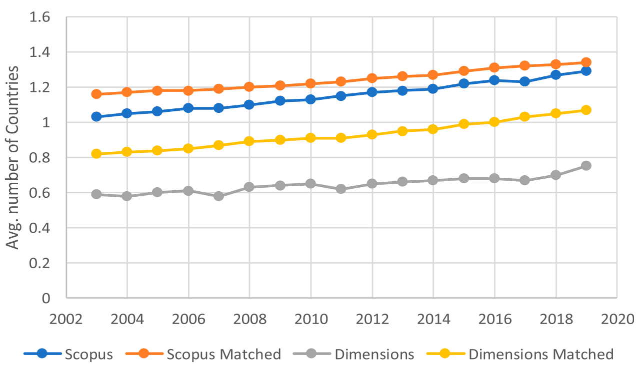

FIGURE 3 | Evolution of the average number of countries per document in Scopus and Dimensions in total and in the matched subsets.

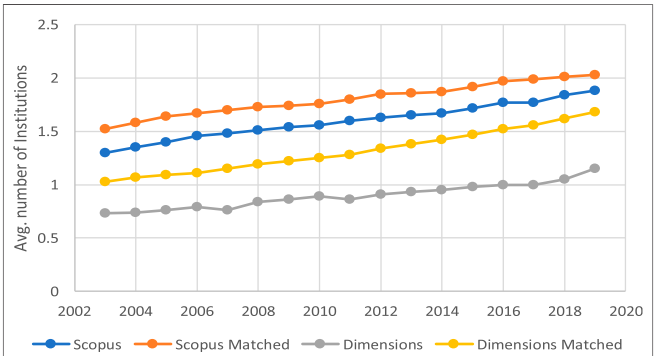

FIGURE 4 | Evolution of the average number of institutions per document in Scopus and Dimensions in total and in the matched subsets.

When the citations of the documents in the two databases are distributed by country, one observes that all of them, regardless of the size of their output, accumulate more citations in the Scopus database than in the Dimensions one. **Figure 8** shows that both total citations and those of matched documents are consistently greater in Scopus than in Dimensions for all countries. The case is similar when the distribution of citations is by institution in the period of observation. The distribution of citations by institution is also greater in Scopus than in Dimensions in more than 97% of the cases. **Figure 9** shows very clearly how just a small group of institutions lies below the straight line, and these conform to the 2.5% of cases that have more citations in Dimensions than in Scopus.

DISCUSSION

Our starting hypothesis was that the difference in overall coverage between the two databases should be similar in general terms when the total set of documents was fragmented into smaller levels of aggregation. From our perspective, it is important that overall coverage levels be maintained on average when the source is split into smaller groupings (countries or institutions, for example) in order to guarantee the bibliometric relevance of the source. For this reason, we continued along the path begun by other workers trying to deepen the comparative analysis of the coverage of the two sources.

Our first conclusion is that, for reasons that have to do with the data structures themselves, the two sources have notable differences in coverage at the level of countries and institutions, with a tendency for there to be greater coverage at those levels in Scopus than in Dimensions. This is even though what was to be expected would have been the

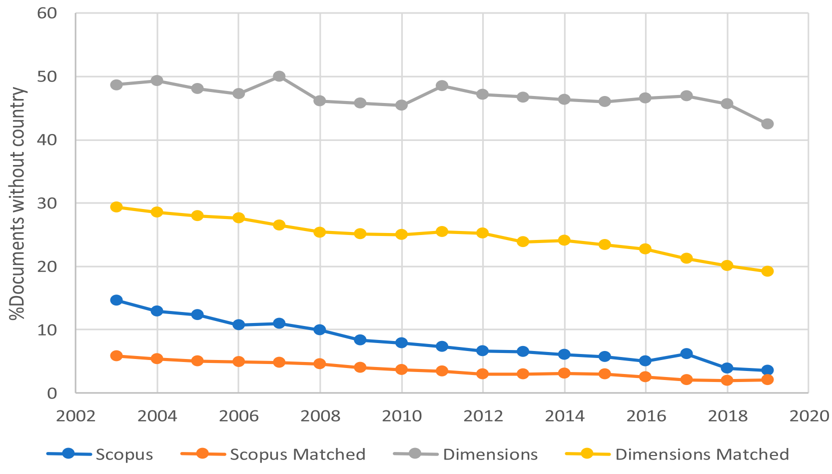

FIGURE 5 | Evolution of the annual percentage of items without country in the four subsets of documents belonging to Dimensions and Scopus.

TABLE 4 | Distribution of document types where no country affiliation data is available.

	Article				**Book**				**Chapter**			
	Total	**Yes**	**No**	**%**	**Total**	**Yes**	**No**	**%**	**Total**	**Yes**	**No**	**%**
2003	1,594,777	847,696	747,081	46.85	5,043	0	5,043	100.00	178,482	60,138	118,344	66.31
2004	1,710,220	921,384	788,836	46.12	5,151	0	5,151	100.00	236,614	65,718	170,896	72.23
2005	1,790,924	969,678	821,246	45.86	6,133	0	6,133	100.00	227,321	75,457	151,864	66.81
2006	1,929,725	1,072,799	856,926	44.41	6,359	0	6,359	100.00	264,507	89,235	175,272	66.26
2007	2,008,313	1,132,294	876,019	43.62	7,561	0	7,561	100.00	471,557	99,410	372,147	78.92
2008	2,108,438	1,193,113	915,325	43.41	7,471	0	7,471	100.00	324,582	112,560	212,022	65.32
2009	2,210,781	1,272,372	938,409	42.45	8,172	0	8,172	100.00	350,024	125,995	224,029	64.00
2010	2,323,835	1,339,765	984,070	42.35	9,405	0	9,405	100.00	312,837	115,127	197,710	63.20
2011	2,555,664	1,483,092	1,072,572	41.97	10,373	0	10,373	100.00	503,972	123,081	380,891	75.58
2012	2,742,694	1,607,802	1,134,892	41.38	12,258	0	12,258	100.00	425,005	127,446	297,559	70.01
2013	2,938,822	1,714,338	1,224,484	41.67	12,181	0	12,181	100.00	474,432	142,900	331,532	69.88
2014	3,122,791	1,811,016	1,311,775	42.01	12,146	0	12,146	100.00	477,231	154,709	322,522	67.58
2015	3,266,544	1,884,432	1,382,112	42.31	13,043	0	13,043	100.00	414,925	154,310	260,615	62.81
2016	3,430,797	1,944,920	1,485,877	43.31	14,272	0	14,272	100.00	377,731	155,480	222,251	58.84
2017	3,652,464	2,076,024	1,576,440	43.16	15,196	0	15,196	100.00	440,965	167,278	273,687	62.07
2018	3,863,842	2,276,994	1,586,848	41.07	17,308	0	17,308	100.00	502,279	182,953	319,326	63.58
Growth rate	142.28	168.61	112.41	−12.33	243.21	0.00	243.21	0.00	181.42	204.22	169.83	−4.12

	Monograph				**Preprint**				**Proceeding**			
	Total	**Yes**	**No**	**%**	**Total**	**Yes**	**No**	**%**	**Total**	**Yes**	**No**	**%**
2003	12,579	1,146	11,433	90.89	48,039	9,309	38,730	80.62	127,949	91,072	36,877	28.82
2004	12,827	1,083	11,744	91.56	50,798	8,041	42,757	84.17	142,125	97,487	44,638	31.41
2005	12,593	506	12,087	95.98	55,872	8,847	47,025	84.17	172,435	121,144	51,291	29.75
2006	12,339	537	11,802	95.65	61,009	9,552	51,457	84.34	198,944	131,960	66,984	33.67
2007	14,005	736	13,269	94.74	68,801	10,669	58,132	84.49	208,261	144,690	63,571	30.52
2008	14,403	875	13,528	93.92	75,226	11,916	63,310	84.16	231,126	168,755	62,371	26.99
2009	18,709	1,066	17,643	94.30	84,053	13,431	70,622	84.02	263,563	177,408	86,155	32.69
2010	20,997	1,429	19,568	93.19	93,239	14,628	78,611	84.31	307,112	204,415	102,697	33.44
2011	21,356	1,719	19,637	91.95	103,214	16,121	87,093	84.38	295,358	172,933	122,425	41.45
2012	28,405	1,777	26,628	93.74	114,152	18,305	95,847	83.96	276,667	147,489	129,178	46.69
2013	39,622	2,484	37,138	93.73	120,577	17,205	103,372	85.73	273,391	178,389	95,002	34.75
2014	33,389	3,073	30,316	90.80	125,340	18,374	106,966	85.34	298,898	196,218	102,680	34.35
2015	31,467	3,346	28,121	89.37	134,199	19,462	114,737	85.50	333,259	201,161	132,098	39.64
2016	38,941	4,478	34,463	88.50	144,704	21,845	122,859	84.90	370,153	211,794	158,359	42.78
2017	41,935	4,854	37,081	88.42	164,527	28,763	135,764	82.52	405,166	227,657	177,509	43.81
2018	37,792	5,292	32,500	86.00	193,204	37,024	156,180	80.84	419,014	232,398	186,616	44.54
Growth rate	200.44	361.78	184.26	−5.38	302.18	297.72	303.25	0.27	227.49	155.18	406.05	54.53

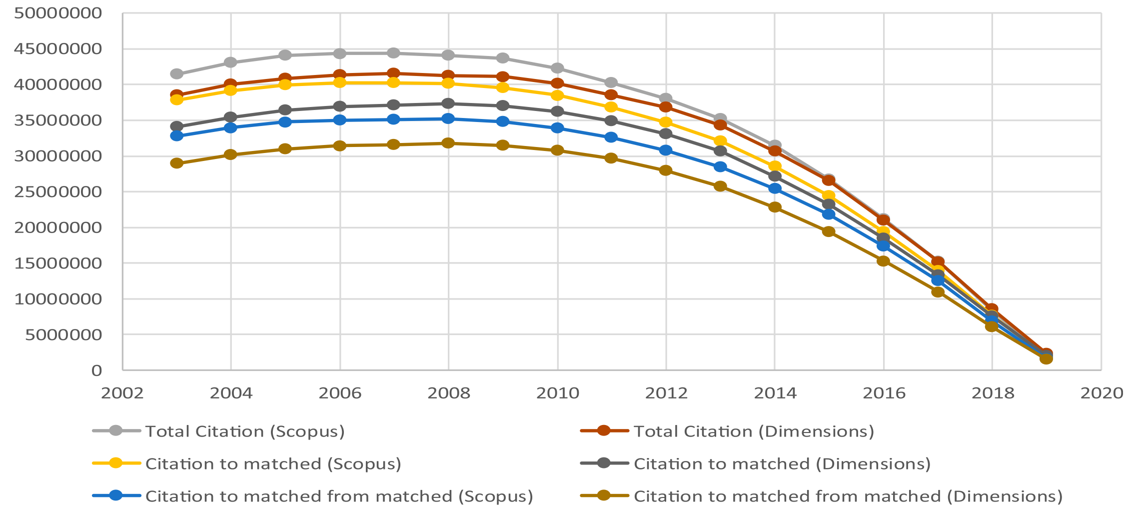

FIGURE 6 | Citations by cited year.

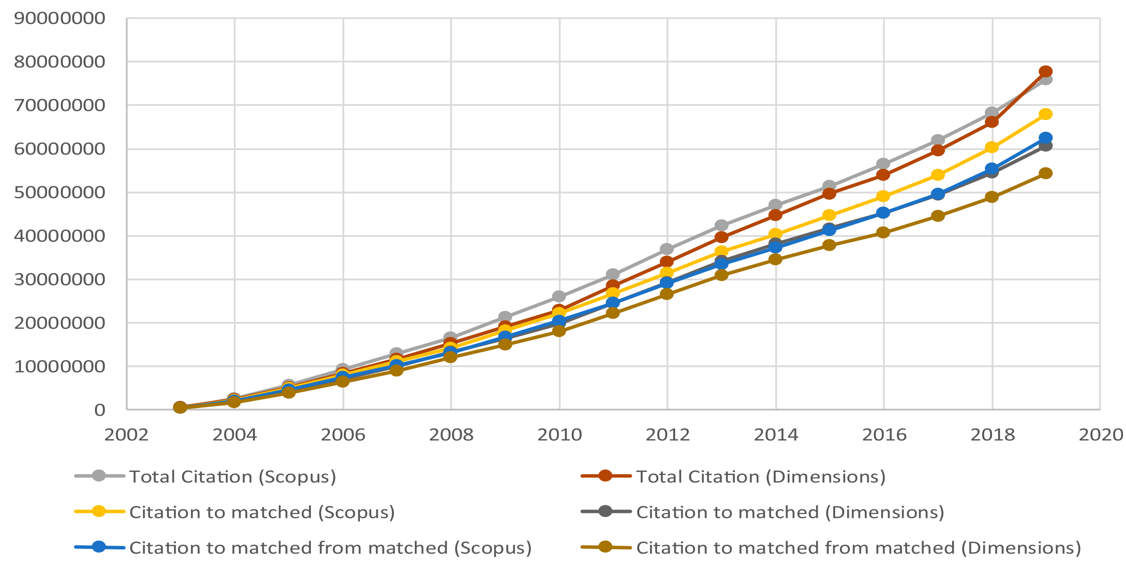

FIGURE 7 | Citations by citing year.

opposite, given the overall differences in coverage between the two sources.

Second, despite the fact that Dimensions has a larger raw coverage of documents than Scopus, close to half of the documents in Dimensions lack country or institutional affiliation information, which means that when documents are aggregated by country or institutional affiliation, Scopus systematically provides more documents/citations than Dimensions. In 2014, Dimensions started working on the problem of creating an entity list for organizations to provide a consistent view of an organization within one content source, but also across the various different types of content. This was the GRID (Global Research Identifier Database) system. At that time, a set of policies about how to handle the definition of a research entity was developed.[2] At the time of writing, GRID contains 98,332 unique organizations, for which the data has been curated and each institution assigned a persistent identifier. This set of institutions represents an international coverage of the world's leading research organizations, indexing 92% of funding allocated globally. It is clear, however, that the repeated differences between Scopus and Dimensions in output and citation are related to the fact that Dimensions' method of linking institutional affiliations to GRID, while a promising idea, is still a work in progress. In overall terms, currently, it limits linkages of item with countries and institutions. This situation mainly affects the possibilities that

[2]https://www.grid.ac/pages/policies

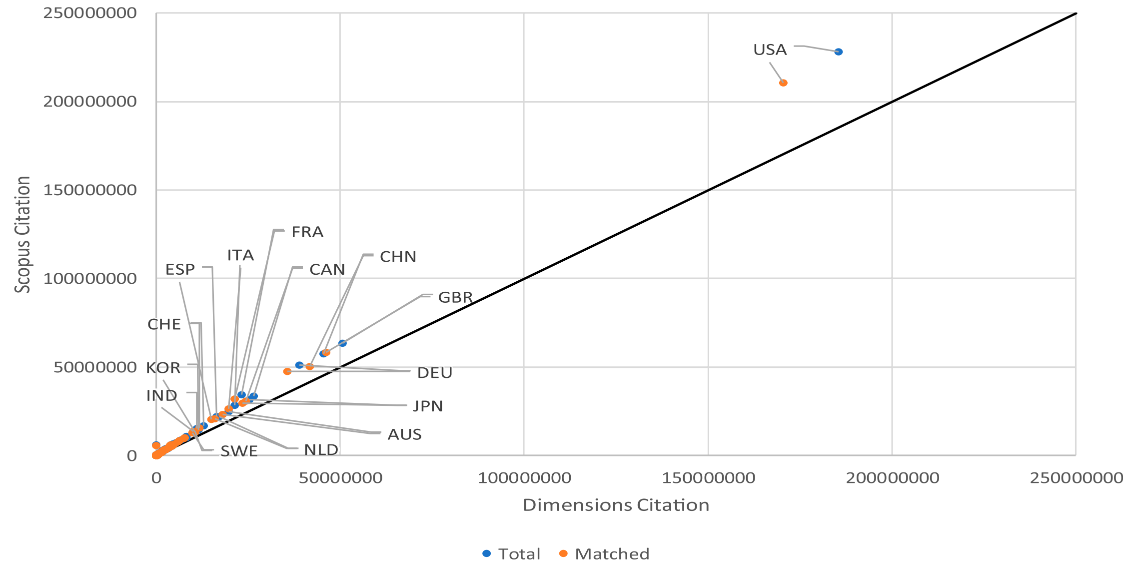

FIGURE 8 | Relationship between total citations and matched documents by country.

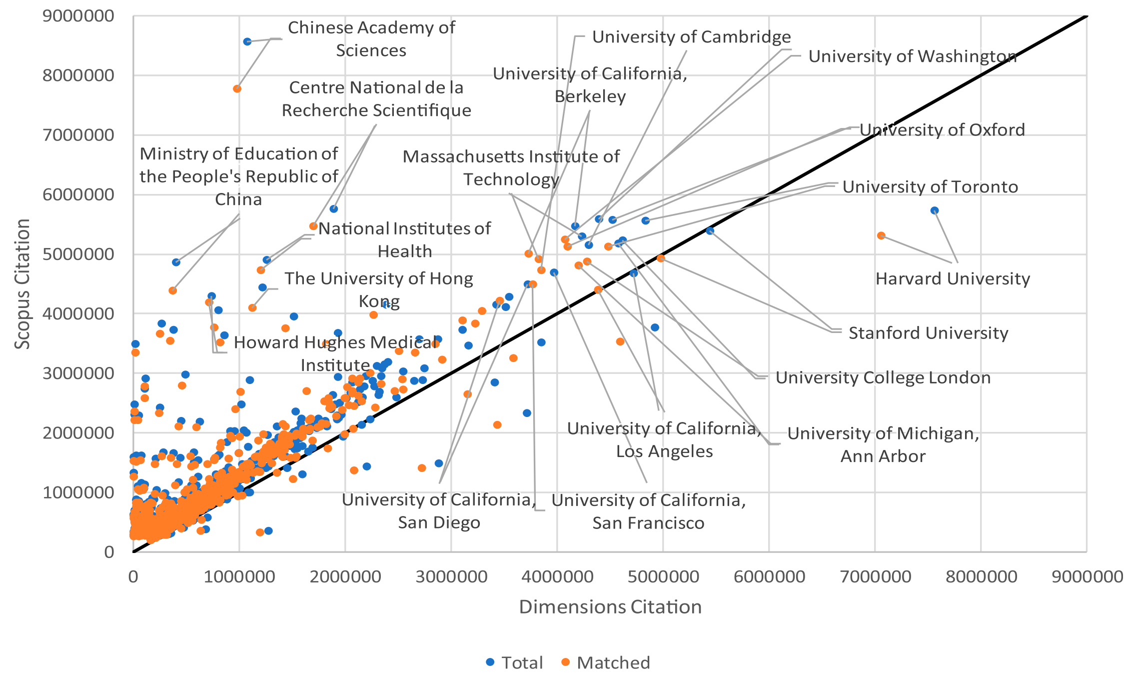

FIGURE 9 | Relationship between total citations and matched documents by institution.

the two sources can offer as instruments for carrying out bibliometric analyses.

As Bode et al. (2019) point out in Dimensions' Guide v.6 (p. 3), "Linked and integrated data from multiple sources are core to Dimensions. These matchings are data driven, then, the content and enrichment pipeline is as automated as possible. However, while an automated approach allows us to offer a more open, free approach it also results in some data issues, which we will continue to have to work on and improve." This is advisable for both the publications and citation links because, as Visser et al. (2020) noted, "Dimensions incorrectly has not identified citation links. Hence, this data source fails to identify a substantial number of citation links" (p. 20). Dimensions also has the limitation that it does not provide data for references that have not been matched with a cited document (p. 23).

The results described should help fill the gap in exploring differences between Scopus and Dimensions at the country and institutional levels. **Figure 5** appears to be the main cause that explains most of the other results. Most of the other results in this manuscript are an effect or consequence of this. This should allow a profile of Dimensions to be outlined in terms of its coverage by different levels of aggregation of its publications in comparison with Scopus. Both of these aspects are highly pragmatic considerations for bibliometric researchers and practitioners, in particular for policymakers who rely on such databases as a principal criterion for research assessment (hiring, promotion, and funding).

At the country level, this study has shown that not all articles had complete address data. Even though there was a decreasing trend over time in the number of documents with no country information in the address data, in 2018 still more than 40% of documents in Dimensions remained without a country. Given the size of the data source and its goal in the scientific market, missing information of the country in the affiliation data has important implications at all levels of aggregation and analysis. Thus, Dimensions does not currently appear to be a reliable data source with which to define and evaluate the set of output at the country level.

At the institutional level, according to Huang et al. (2020), "Universities are increasingly evaluated on the basis of their outputs which are often converted to rankings with substantial implications for recruitment, income, and perceived prestige." The present study has shown that Dimensions does not record all institutional affiliation of the authors, which has implications for metrics and rankings at the institutional scale. In this case, it seems advisable to integrate diverse data sources into any institutional evaluation framework (Huang et al., 2020).

We have not been comparing document types but presenting results derived from the matching procedure. As in Visser et al. (2020), we found that there were many articles in Dimensions for which there was no matching document in our matching procedure. This is because it seems that any document published in a journal is classified as an article in Dimensions.

Finally, as in previous studies examining data sources' coverage (Moya-Anegón et al., 2007), to very briefly conclude and with possible future bibliometric studies in mind, the above considerations conform to an important part of the context of scientific output and evaluation and should be taken into account so as to avoid bias in the comparison of research results in diverse domains or at different aggregation levels. All data sources suffer from problems of incompleteness and inaccuracy of citation links (Visser et al., 2020, p. 23), and GRID is not yet perfect and never will be (Bode et al., 2019, p. 6). But we are confident that studies like the present will help to improve this tool and the data in the near future.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because the SCImago group annually receives a raw data copy in XML format through a contract with Elsevier. The SCImago group has the possibility of downloading a copy of Dimensions in Json format through an agreement with Digital Science. We are not allowed to redistribute the Scopus and Dimensions data used in this paper. Requests to access the datasets should be directed to felix.moya@scimago.es.

AUTHOR CONTRIBUTIONS

VG-B: conception, data curation, and writing. AM: data curation. ZC-R: conception, data analysis, and writing. FM-A: conception, data analysis, and writing. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

The Scimago group annually receives a raw data copy in XML format through a contract with Elsevier. The Scimago group has the possibility of downloading a copy of Dimensions in JSON format through an agreement with Digital Science. We are not allowed to redistribute the Scopus and Dimensions data used in this paper.We thank the Dimensions and Scopus development teams for their availability to provide us with access to the information necessary to carry out this analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frma.2020.593494/full#supplementary-material.

REFERENCES

Adams, J., Jones, P., Porter, S., Szomszor, M., Draux, H., and Osipov, I. (2018). Dimensions–A collaborative approach to enhancing research discovery. Technical report. Digital Science. doi:10.6084/m9.figshare.5783160.v1

Archambault, É., Campbell, D., Gingras, Y., and Larivière, V. (2009). Comparing bibliometric statistics obtained from the Web of Science and Scopus. *J. Am. Soc. Inf. Sci.* 60 (7), 1320–1326. doi:10.1002/asi.21062

Baas, J., Schotten, M., Plume, A., Côté, G., and Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quant. Sci. Stud.* 1 (1), 377–386. doi:10.1162/qss_a_00019

Bode, C., Herzog, C., Hook, D., and McGrath, R. (2019). A guide to the dimensions data approach. Technical report. Digital Science. doi:10.6084/m9.figshare.5783094.v5

Bornmann, L. (2018). Field classification of publications in Dimensions: a first case study testing its reliability and validity. *Scientometrics* 117 (1), 637–640. doi:10.1007/s11192-018-2855-y

Gorraiz, J., Gumpenberger, C., and Wieland, M. (2011). Galton 2011 revisited: a bibliometric journey in the footprints of a universal genius. *Scientometrics*, 88 (2), 627–652. doi:10.1007/s11192-011-0393-y

Guerrero-Bote, V. P., and Moya-Anegón, F. (2015). Analysis of scientific production in food science from 2003 to 2013. *J. Food Sci.* 80 (12), R2619–R2626. doi:10.1111/1750-3841.13108

Guerrero-Bote, V. P., Sánchez-Jiménez, R., and Moya-Anegón, F. (2019). The citation from patents to scientific output revisited: a new approach to Patstat/Scopus matching. *El Prof. Inf.* 28 (4), e280401. doi:10.3145/epi.2019.jul.01

Hane, P. (2004). Elsevier announces Scopus service. *Information Today*. Available at: http://newsbreaks.infotoday.com/nbreader.asp?ArticleID=16494 (Accessed June 9, 2011).

Harzing, A.-W. (2019). Two new kids on the block: how do crossref and dimensions compare with Google scholar, Microsoft academic, Scopus and the Web of Science? *Scientometrics*, 120 (1), 341–349. doi:10.1007/s11192-019-03114-y

Herzog, C., Hook, D., and Konkiel, S. (2020). Dimensions: bringing down barriers between scientometricians and data. *Quant. Sci Stud.* 1 (1), 387–395. doi:10.1162/qss_a_00020

Hook, D. W., Porter, S. J., and Herzog, C. (2018). Dimensions: building context for search and evaluation. *Front. Res. Metr. Anal.* 3, 23. doi:10.3389/frma.2018.00023

Huang, C.-K., Neylon, C., Brookes-Kenworthy, C., Hosking, R., Montgomery, L., Wilson, K., et al. (2020). Comparison of bibliographic data sources: implications for the robustness of university rankings. *Quant. Sci. Stud.* 1, 445–478. doi:10.1162/qss_a_00031

Jacsó, P. (2011). The h-index, h-core citation rate and the bibliometric profile of the Scopus database. *Online Inf. Rev.* 35 (3), 492–501. doi:10.1108/14684521111151487

Leydesdorff, L., Moya Anegón, F., and Guerrero Bote, V. P. (2010). Journal maps on the basis of Scopus data: a comparison with the journal citation reports of the ISI. *J. Am. Soc. Inf. Sci. Technol.* 61 (2), 352–369. doi:10.1002/asi.21250

Martín-Martín, A., Orduna-Malea, E., and Delgado López-Cózar, E. (2018a). Coverage of highly-cited documents in Google scholar, Web of Science, and Scopus: a multidisciplinary comparison. *Scientometrics* 116 (3), 2175–2188. doi:10.1007/s11192-018-2820-9

Martín-Martín, A., Orduna-Malea, E., Thelwall, M., and Delgado López-Cózar, E. (2018b). Google Scholar, Web of Science, and Scopus: a systematic comparison of citations in 252 subject categories. *J. Inf.* 12 (4), 1160–1177. doi:10.1016/j.joi.2018.09.002

Martín-Martín, A., Thelwall, M., Orduna-Malea, E., and López-Cózar, E. D. (2020). Google scholar, Microsoft academic, Scopus, dimensions, Web of Science, and OpenCitations' COCI: a multidisciplinary comparison of coverage via citations. arXiv:2004.14329 [Preprint].

Mongeon, P., and Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics* 106 (1), 213–228. doi:10.1007/s11192-015-1765-5

Moya-Anegón, F., Chinchilla-Rodríguez, Z., Vargas-Quesada, B., Corera-Álvarez, E., González-Molina, A., Muñoz-Fernández, F. J., et al. (2007). Coverage analysis of SCOPUS: a journal metric approach. *Scientometrics* 73 (1), 57–58. doi:10.1007/s11192-007-1681-4

Moya-Anegón, F., Guerrero-Bote, V. P., Lopez-Illescas, C., and Moed, H. F. (2018). Statistical relationships between corresponding authorship, international co-authorship and citation impact of national research systems. *J. Inf.* 12 (4), 1251–1262. doi:10.1016/j.joi.2018.10.004

Pickering, B. (2004). Elsevier prepares Scopus to rival ISI Web of Science. Information World Review, 200, 1.

Thelwall, M. (2018). Dimensions: a competitor to Scopus and the Web of Science?. *J. Inf.* 12, 430–435. doi:10.1016/j.joi.2018.03.006

Visser, M., van Eck, N., and Waltman, L. (2020). Large-scale comparison of bibliographic data sources: Scopus, Web of Science, dimensions, crossref, and Microsoft academic. arXiv:2005.10732 [Preprint]. Available at: https://arxiv.org/abs/2005.10732

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Research Metrics and Analytics

CONCEPTUAL ANALYSIS
published: 25 January 2021
doi: 10.3389/frma.2020.614016

In Defense of Merit to Overcome Merit

Cinzia Daraio *

Department of Computer, Control and Management Engineering A. Ruberti (DIAG), Sapienza University of Rome, Rome, Italy

Bibliometric indicators such as the number of published articles and citations received are subject to a strong ambiguity. A high numerical value of bibliometric indicators may not measure the quality of scientific production, but only a high level of activity of a researcher. There may be cases of good researchers who do not produce a high number of articles, but have few research products of high quality. The sociology of science relies on the so-called "Matthew effect," which is inspired by Matthew's Gospel on Talents. "Those that have more will have more" seems to support the idea that those that publish more, merit to have higher bibliometric indicators, and to be recognized for their major results. But is this really the case? Can bibliometric indicators be considered a measure of the merit of scholars or they come from luck and chance? The answer is of fundamental importance to identify best practices in research assessment. In this work, using philosophical argumentation, we show how Christian theology, in particular St. Thomas Aquinas, can help us to clarify the concept of merit, overcoming the conceptual ambiguities and problems highlighted by the existing literature. By doing this, Christian theology, will allow us to introduce the evaluation framework in a broader perspective better suited to the interpretation of the complexity of research evaluation.

Keywords: research assessment, bibliometrics, best-practices, Christian theology, St Thomas Aquinas

Edited by:
Juan Ignacio Gorraiz, Universität Wien, Austria

Reviewed by:
Steve Reding, Medical University of Vienna, Austria
Nicola De Bellis, University of Modena and Reggio Emilia, Italy

***Correspondence:**
Cinzia Daraio
daraio@diag.uniroma1.it

Specialty section:
This article was submitted to Research Assessment, a section of the journal Frontiers in Research Metrics and Analytics

Received: *04 October 2020*
Accepted: *09 December 2020*
Published: *25 January 2021*

Citation:
Daraio C (2021) In Defense of Merit to Overcome Merit. Front. Res. Metr. Anal. 5:614016. doi: 10.3389/frma.2020.614016

INTRODUCTION

The scientific productivity of researchers follows quantitative rules known since the last century. The law introduced by Lotka (1926) is well known: It is an inverse-square law of productivity according to which the number of people producing n papers is proportional to $1/n^2$. For every 100 authors who produce a single paper in a certain period, there are 25 with two, 11 with three, and so on. Another well-known law is the Price's law (De Solla Price, 1963) which refers to the relationship between the literature on a subject and the number of authors in the subject area, and states that half of the publications come from the square root of all contributors. These empirical laws can be linked to the so-called "Matthew effect," based on the Parable of the Talents (see **Supplementary Appendix 1** for the full text)" on which the sociology of science (Merton, 1973) developed. A rich literature has analyzed the skewness of scientific productivity distributions (Seglen, 1992) across the sciences (Albarrán et al., 2011; Ruiz-Castillo and Costas, 2014) and has investigated the connected cumulative advantages (Allison and Stewart, 1974) and the related inequalities (Allison, 1980; Allison et al., 1982).

The bodies of literature cited above show the intrinsic inequality of scientific productivity, a sort of *undemocratic* nature inherent in scientific production/productivity, as De Solla Price nicely illustrated in his famous 1963 book *Little science, big science… and beyond:*

> "About this process there is the same sort of essential, built-in undemocracy that gives us a nation of cities rather than a country steadily approximating a state of uniform population

> density. Scientists tend to congregate in fields, in institutions, in countries, and in the use of certain journals. They do not spread out uniformly, however desirable that may or may not be. In particular, the growth is such as to keep relatively constant the balance between the few giants and the mass of pygmies. The number of giants grows so much more slowly than the entire population that there must be more and more pygmies per giant, deploring their own lack of stature and wondering why it is that neither man nor nature pushes us toward egalitarian uniformity (de Solla Price, 1963, p. 59)."

Xie (2014) distinguishes three kind of inequalities across scientists: resources, research outcomes, and rewards (monetary or nonmonetary) each of which is influenced by the institutional and country contexts. The ongoing trend toward increasing quantitative assessments (based on bibliometric indicators), which amplifies the Matthew effect, produces an exacerbation of inequalities in science.

The high inequality in scientific rewards is often defended on the ground of both the positive externalities generated by science and the merit-based evaluation in place (see more discussion in Xie, 2014).

It is important to distinguish among different but interconnected activities such as research evaluation, reward distribution and research management. Different forms (e.g. individual/disciplinary orientated vs. collective/policy orientated evaluation) and fora (e.g. hiring committees vs. institutional evaluation) in which research assessment and evaluation takes place, exist. The evaluation of interest to a national evaluation agency such as the Italian ANVUR, aiming to the most productive allocation of resources is different from the evaluation which affects single institutions or even individuals, in which the consideration of the right evaluation in interpersonal terms, is important.

In this paper we do not intend to analyze research evaluation, reward distribution and research management in detail. We aim at addressing the ambiguity and content of bibliometric indicators taking one step back and reflecting on the multiple meanings that underlay a concept or rationale like merit that is omnipresent in the realm of bibliometrics/scientometrics and research evaluation. By doing this, we will offer a wider framework for assessing research which will eventually be useful for better characterizing the distribution of rewards and the management of research.

Bibliometric indicators such as the number of published articles or citations received are currently used in evaluation exercises for hiring new scholars and/or to promote researchers of universities or research centers. One reason for their success is the availability of standardized data and information and their simplicity. The very existence and success of evaluative bibliometrics depends indeed on the (possibly utopian) search for non-subjective and non-individual-related traces of epistemic value. On the other hand, papers and citations can certainly be very misleading indicators of scientific achievement. Among the criticisms addressed to the use of bibliometric indicators we find their *inability to discriminate* between high-quality scientific contributions and mere volumes of scientific production, and the *unintended consequences* generated by their use on the behavior of scientists (see e.g. Dahler-Larsen, 2014; De Rijcke et al., 2016; Biagioli and Lippman, 2020).

Using a straightforward model, we argued (Ruocco, Daraio et al., 2017) that the distributions of the individual bibliometric indicators observed might be the result of chance and noise (chaos) related to multiplicative phenomena connected to a *publish or perish* inflationary mechanism, led by scholars' recognition and reputations. This interpretation leads us to cast some doubts on the use of the number of papers and/or citations as a measure of scientific achievements. In the conclusion we wrote:

> "A tricky issue seems to emerge from this interpretation of our model that is: what do bibliometric indicators really measure? The analysis of this issue calls for deeper investigations on the meaning of the bibliometric indicators. These further analyses are clearly outside the purpose of the present paper. They will require the development of more detailed and accurate models than our (over)simplified model, in which the relationships among intelligence, talents, their historical characterization, ability, merits and their measure are more carefully taken into account and modelled. This is an interesting and intriguing topic for further research to be carried out beyond Science of Science and Sociology of Science, including elements and investigation tools from Philosophy, Psychology and Theology. (Ruocco, Daraio et al., 2017, p. 7, p. 7)."

In this work, we carry on this line of research, trying to tackle the problem of the content and ambiguity of bibliometric indicators, which is very relevant for the evaluation of performance and the identification of best practices.

AIM AND CONTRIBUTION

The dialectical method of Scholasticism and its rediscovery of the Aristotelian way of dealing analytically with empirical questions have arguably played a positive role in Western intellectual development. This especially as western universities (and obviously global universities orientated toward the ideal of western universities) until today need to be understood as steeped in a tradition of Christian philosophy and thus largely drawing back on reasoning (-s) (also) emerging from theology.

The aim of this work is to address the ambiguity of bibliometric indicators, that is, of what they measure and in particular whether they measure the merits of researchers rather than luck or chance, starting with the clarification of the concept of merit. Using philosophical argumentation, we attempt to show the usefulness of Christian theology, or the science of faith, to clarify the concept of merit, overcoming the conceptual ambiguities and problems highlighted by the existing literature, as rightly emphasized by Sen (2000):

> "The idea of meritocracy may have many virtues, but clarity is not one of them. [. . .] Meritocracy, and more generally the practice of rewarding merit, is essentially underdefined, and we cannot be sure about its content—and thus about the claims regarding its "justice"—until some further specifications are made (concerning, in particular, the objectives to be pursued, in terms of which merit is to be, ultimately, judged). The merit of actions—and (derivatively) that of persons performing actions—cannot be judged independent of the way we understand the nature of a good (or an acceptable) society" (Sen, 2000; 6–7).

Our purpose is to show that Christian theology, in particular through the thought of St. Thomas Aquinas, may allow us to clarify the concept of merit and connect merit to other related concepts, putting the evaluation framework in a broader perspective which enables it to deal with the complexity of research evaluation.

RELATED WORKS

Leaving aside for the moment the conceptual complexity of merit, we will consider a starting definition of merit as "just compensation" and will consider meritocracy as a "system of evaluation and enhancement of individuals, based exclusively on the recognition of their merit." The existing literature on merit is very rich and often based on ideological positions in favor or against merit. In the practice of research evaluation, often those in favor of the use of bibliometric indicators also support merit and the application of meritocracy, while those who are against evaluation in general are also opposed to merit and the application of meritocratic evaluations, often using as arguments, the difficulty or impossibility of measuring merit in the scientific field.

Meritocracy has come under increasing criticism in recent years. There is a rich and growing literature against merit (Young, 1958, 1994; Bell, 1972; Daniels, 1978; Arrow et al., 2000; Brown, 2001; Castilla and Benard, 2010; McNamee and Miller, 2014; Frank, 2016; Littler, 2018; Mijs, 2019; Sandel, 2020), just as the literature against evaluation is dense and growing (Abelhauser et al., 2011; Del Rey, 2013; Berg and Seeber, 2016; Gingras, 2016; Muller, 2018).

The book "The Rise of Meritocracy" (Young, 1958) introduced the term meritocracy in a negative way, showing a dystopian future in which an emergent elitism of meritocratic people selected on the base of their merits assessed through the evaluation of their intelligence and efforts, without considering other factors (such as ethnicity and gender) reinforce the *status quo* favoring dominant groups that control the evaluation process (Young, 1994). Conditions such as inheritance, social advantages, and discrimination that may hamper accurate merit-based outcome allocations are usually neglected (McNamee and Miller, 2004). Therefore, meritocracy is ideologically considered as a form of hegemony which consolidates and legitimizes social inequality. Littler (2018, p. 3-12) summarizes the five problems of meritocracy listed below.

(1) The first issue relates to the consideration that meritocracy endorses a competitive and hierarchical system which legitimizes inequality and damages community advancing self-interest and highly competitive people.
(2) The second issue is connected to the assumption that talent and intelligence are typically innate: they depend on an essentialized conception of intellect and aptitude.
(3) The third issue of meritocracy is that it does not consider the impact of different contexts. Social, institutional and national contextual differences can strongly affect performance.
(4) The fourth issue is the uncritical support of meritocracy to the current hierarchy of professions, endorsing the status quo. Related to this issue, Castilla and Benard (2010) show that in the managerial profession, when the organizational culture explicitly promotes meritocracy, there is a greater bias in favor of men over equally performing women, and call this as the "paradox of meritocracy."
(5) The fifth issue relates the function of meritocracy as an "ideological myth" to hide and amplify economic and social inequalities. This last point is discussed by many other studies. For instance, in the book "The Meritocracy Myth" by McNamee and Miller (2014) the authors about the connection of merit with social inequality state:

> "Currently in the United States inequality is "legitimized," or "explained," predominately by an ideology of meritocracy. America is seen as the land of opportunity where people get out of the system what they put into it. Ostensibly, the most talented, hardest working, and most virtuous get ahead. The lazy, shiftless, and inept fall behind. In this formulation, you may not be held responsible for where you start out in life, but you are responsible for where you end up because the system is "fair" and provides ample opportunity to get ahead. An important aspect of ideologies of inequality is that they do not have to be objectively "true" to persuade those who have less to accept less (McNamee and Miller, 2014, p. 3, p. 3)."

Along the same line, Bell (1972) and Mijs (2019) state that citizens' approval to inequality is explained by their persuasion that the success of society reflexes a meritocratic process. The rising inequality is legitimated by the popular credence that the income gap is meritocratically deserved.

Arrow et al. (2000) analyze deeply economic inequality and their connection with meritocracy investigating the interconnections among merit, reward and opportunity; causes and consequences of intelligence; schooling and economic opportunity and policy options; in Brown (2001) the interested readers can find a comprehensive review of its content.

Finally, Sandel (2020) describes the problems generated by meritocracy among the winners and the harsh judgment it imposes on those left behind. He offers an alternative way of thinking about success, more attentive to the role of luck in human activities, more helpful to an ethic of humility, and more open to a politics of the common good.

On the other hand, the social psychological literature has long conceptualized meritocracy as a principle of distributive justice. Sen (2000) states that rewarding "merit" must provide incentives that contribute to social welfare. Meritocracy is viewed positively in Miller (1996), Heneman (2002) and Son Hing et al. (2011). Meritocracy is invoked in all recent teaching and research evaluation regulations and laws, and it is positively considered both by researchers and by the public opinion, at least in Italy. The Italian law n. 240 of 2010 on "Regulations on the organization of universities, academic staff and recruitment, as well as delegation to the Government to encourage the quality and efficiency of the university system," called the Gelmini Law, cites the merit more than ten times, saying that "the ministry (...) enhances merit," "the ministry [...] verifies and evaluates the results according to criteria of quality, transparency and promotion of merit," "allocation of resources among universities and selection of the recipients of the intervention according to academic and scientific merit criteria ". The criterion of merit is present in the ethical codes of all Italian universities. In that of the University of Rome La Sapienza, for example, it is reported: "the Code commits all members of the academic community to adopt behaviors suitable for: "e) pursuing and guaranteeing compliance with the merit criterion in all circumstances, taking into account, when possible, the indicators used in international scientific teaching community. " The criterion of merit is also considered for the scientific evaluation of European projects, as stated e.g. in the documents European Commission, 2013, Ethics for researchers Facilitating Research Excellence in FP7 and in the European Commission, H2020 Ethics Manual.

In particular, Heneman (2002) supports meritocracy when it is issued for the right reasons and attention is paid to strategy and implementation questions; in these cases, merit can be a viable reward program. Jones (1994) exposes his support for a meritocratic system in the management of firms: "under certain circumstances managers are morally justified in making personal decisions based solely on merit." Simon (1974) tried to bring out the moral foundation of the meritorian principle, identifying the conflictual relationships between merit, equality and "gifts" or natural talents received and not connected to our efforts, and points out that it is not possible to dismiss the merit and the need to conjugate distributive justice, meritocratic distribution with compensatory justice.

Young (1958) defines merit as the sum of intelligence and effort. Nevertheless, one of the primary concerns with meritocracy is the ambiguous (unclear) definition of "merit" (Arrow et al., 2000; Sen, 2000). Carson in his book of 2007 *The Measure of Merit* shows that talents and intelligence have become constituents of the societies in which they were produced and adopted, continually shaping and being shaped by these cultures. The concepts of intelligence and merit, hence, remain always contestable terms in the recurrent debates about the social and political implications of inequality for a modern democracy (Carson, 2007). In addition, from a history of quantification perspective, recently, Carson (2020) points out that

> "quantification and measurement should be seen not just as technical pursuits, but also as normative ones. Every act of seeing, whether through sight or numbers, is also an act of occlusion, of not-seeing. And every move to make decisions more orderly and rational by translating a question into numerical comparisons is also a move to render irrelevant and often invisible the factors that were not included (Carson, 2020, p. 1)."

Other studies (Sternberg and Kaufman, 2011; Kaufman, 2013) show that "greatness" is more than just the sum of the "nature" and "nurture" components, and to understand it we have to go beyond talent and practice. On top of that, there is a literature on the need of evaluation to assess merit, provide incentive and good practice in the assessment including a learning dimension (Nielsen and Hunter, 2013; Vidaillet, 2013).

MATERIALS AND METHODS

From what we have discussed in the previous sections, the clarification of the meaning of "merit" seems then an ineludible step toward our understanding of the role of bibliometric indicators and of what they can measure.

Vanzini (2019) using and updating the thought of St. Thomas, the great philosopher and theologian of the Middle Ages, shows how theology, the science of faith or the knowledge of the Christian faith is grounded on a rational, rigorous and well-founded basis. Theology, therefore, in the vision of Thomas, reveals itself in all its rigor as a scientific discipline. Thomas Aquinas undoubtedly represents one of the most important and influential thinkers in the entire history of Western thought. Some recent researches in the field of the history of medieval philosophy have historically reconstructed his philosophical thought in its entirety, showing its value and relevance (see Porro, 2012). We will use St. Thomas Aquinas thought to shed some light on the complex and ambiguous concept of merit.

Exegesis of the Two Parables of the Gospel of Matthew

Let us start with two parables of the Gospel according to Matthew that apparently show the contradiction and ambiguity of the concept of merit: The parable of the vineyard workers (Mt, 20:1-16) and The Parable of the Talents (Mt, 25: 14-30).

The parable of the workers is the most "scandalous", while the parable of talents is better known to those involved in the evaluation of research. For the convenience of the reader, the full text of the two parables is reported in **Supplementary Appendix 1**. In the parable of the vineyard workers, the landowner of a vineyard hires for a day's work. He hires a few at the first hour of the day, and the salary agreed for a full day's work is one denarius. Then the landowner calls other workers at all hours of the day, even an hour before the end of the day. With the newly called, the landowner does not agree on a precise wage, but simply says: "I'll give you whatever is right." To the workers of the last hour he does not even say this. The parable leads the listener to ask himself: how will the landowner behave with the latter? The answer is confusing, completely unexpected: the landowner gives everyone the same pay, even the last ones. It is not fair, say the workers of the first hour. And certainly the readers think the same thing: a single hour of work does not deserve the same wage as a whole day. This is a complex parable: a complete analysis of its text is beyond the scope of this work (see Maggioni, 2009)). We will focus only on the paradox of the landowner's injustice, to try to understand why he gives everyone, even to the last hour workers, the same wage as the former? Is it a form of injustice? And what kind of "merit" does he apply? St. Thomas's commentary on the Gospel of Matthew (recently published in Italian by Edizioni Studio Domenicano in 2018, see D'Aquino, 2018) will help us to clarify this issue. St. Thomas's commentary on verses 13–15 (see D'Aquino, 2018, p. 339) explains the logic of this paradox i.e. the apparent landowner's injustice. St. Thomas comments stating that first he shows his justice, and his mercy; second the fairness of remuneration. On the first point (his justice and his mercy) he does three things. First, he denies injustice; second he induces the contract, third he induces the remuneration made. He then places the mercy exercised ("I want to give it also to the latter as to you") and the right to exercise it. It is not a question of injustice but rather of the proclamation of God's mercy, of grace. The focus of the parable, for our purpose, is in verse 10 ("So when the first ones came, they assumed they would get more, but they also received a denarius each") and it is clarified by the criticisms that the workers move to the landowner (vv 11–12) and by the reply of the landowner to them (vv 13–15). On closer inspection, the workers of the first hour do not complain about the damage they have suffered (they have agreed on a denarius and received it) but rather for an advantage granted to others. They are envious that others have been treated like them. The wrong they think they suffer is in seeing that the landowner is good to others. It is the envy of the just toward the sinner. Then the Parable could be addressed precisely to the righteous to teach them how to behave in the face of God's mercy.

St. Thomas organizes the Gospel of Matthew into three parts. The parable of the workers in the vineyard is in the second part, including the doctrine of Christ and the end to which it leads. While the parable of the talents is in the third part, in the section on the final judgment. The parable of the talents tells of someone who is excluded from the Kingdom of Heaven (from salvation) because he has not multiplied the goods received. The parable tells of a man who, leaving on a long journey, entrusted his goods to his servants: to one he gave five talents, to one two talents and to another one a talent. To each according to their abilities. After a long time, the master returned and settled the accounts with his servants. The servants who had five and two talents multiplied them and returned ten and four respectively to the master. The master praized them and invited them to enter into his joy. The servant who received only one talent, on the other hand, hid the talent under the ground, and then returned it to the master. The master ordered to take away the talent and give it to the servant with the 10 talents ("whoever has will be given, and whoever does not have, even what he seems to have will be taken away") and he ordered the useless servant to be thrown into outer darkness.

In this parable about talents, we focus on the behaviour of the third servant. The first two servants seem to highlight, by contrast, the behaviour of the third one. Unlike the first two who invest the talents received, the third servant hides his talent in a hole. The focus of the parable, for our purpose, is the dialogue between the wicked servant and the master (vv 24–27). Even the listener in this parable is tempted to hold the reasoning of the wicked servant right and the master's claim unjust. We could say that this reaction is very similar to what we said above about the first hour workers. The conduct of God is not understood; he is considered unjust. Justice is conceived as a mere (simple) relationship of equality.

Justice in the Thought of St. Thomas

St. Thomas defines justice as "the firm and constant will to give each one what is due to him (ST, II-II q.58, a. 1)". As described in Mondin (2000), p. 322), justice for St. Thomas is the virtue that orders man to another and that means that man must always respect this *otherness* because every man is another, a person. The other (each) also embraces the community. Therefore, the indication "to give each his own" contemplates both the duty of the individual to contribute to the common good, and the duty of the community to give its own to individual citizens. St. Thomas, like Aristotle, distinguishes three main forms of justice, namely: distributive, commutative and legal. Distributive justice concerns the duties of the community toward individuals. In *distributive* justice, the burden of giving each his own belongs to the state in relation to the citizens. Commutative justice concerns the duties of justice between private persons. In *commutative* justice, the burden of giving each his own falls to the citizens in mutual relations. Legal justice is about the duties of individuals to the community. In *legal* justice, the burden falls on citizens to the state and consists in observing its laws.

As noted in Mondin (2000, p. 323), all three types of justice studied by St. Thomas belong to *social justice*, even if St. Thomas does not mention the notion of social justice explicitly. It is always a question of duty toward others while safeguarding a certain equality of relationships. Social justice therefore does not nullify the requirements of the three forms of justice but pushes toward their more appropriate and complete application. It points to a *superior model of equity*, which establishes the rights of others, even more than on the consideration of what is strictly due to them on a quantitative level, on the basis of the needs that arise from their dignity as human persons. So naturally, it gives to each his own, according to established legal justice, commutative and distributive, but starts from the recognition of the inalienable rights proper to each person, it has to help in their success and their development. The social justice of St. Thomas, the just price, the just means to calm down, distributive justice combine with other forms of justice and also includes mercy/grace. Merit seems seen from the overall social point of view.

All the thought of St. Thomas is oriented on the principle of equivalence which is the basis and substance of justice. We can see this in the following texts of the *Summa Theologie* (D'Aquino, 2014):

> "If one were to receive something for public services, one would proceed not according to distributive justice, but according to the commutation. In fact, in distributive justice the equivalence between what one receives and what he himself had given is not considered, but the comparison is with what others receive according to their respective conditions". (ST, II-II, q. 61 a. 2, our own translation from the Italian version reported in **Supplementary Appendix 2**).
>
> "[. . .] Ambrose says: "Justice is that virtue which gives each his own, which does not demand the other and which sacrifices his own advantage for the common good.
>
> [. . .] Solution of the difficulties: 1 Since justice is a cardinal virtue, it is accompanied by other secondary virtues, such as mercy, liberality and other virtues of kind, which we will talk about later. Therefore, helping the needy, which belongs to piety or mercy, and benefiting with munificence, which belongs to liberality, are attributed by reduction to justice as to the principal virtue." (ST, II-II, q. 58, a. 11, our own translation from the Italian version reported in **Supplementary Appendix 2**).

In these texts St. Thomas shows us that he does not have a narrow vision of justice even if for him it is correct to say that it consists in giving each his own. Justice appears to have a broader meaning in the *Summa Theologiae*. Also the concept of merit which in the *Summa* is reported under the grace. A concept somehow linked to merit, intended as just compensation, is that of just price according to St. Thomas, including those of just wage, as wage is the price of a particular factor of production (labor).

Some brief but clear passages are present in the *Summa Theologiae,* in the questions 58 and 77:

> "[. . .] 3. As the Philosopher [Aristotle, TN] notes, anything superfluous in matters of justice by extension is called *profit*, and any impairment is called *damage*. And this is because justice is exercised first of all and more universally in the voluntary exchanges of goods, that is, in the sales to which this nomenclature is suitable in the proper sense, and from them it then extends to everything that can be the object of justice. And the same is true for the expression: to give each his own." (ST, II-II, q. 58, a. 11, our own translation from the Italian version reported in **Supplementary Appendix 2**).
>
> "The just price is often not precisely determined, but must be calculated with a certain elasticity, so that small increases or impairments do not compromise the equality of justice." (ST, II-II, q. 77 a. 1, our own translation from the Italian version reported in **Supplementary Appendix 2**).

In the same question 77, we find out a definition of sale:

> "The sale [in itself, TN] was introduced for the common advantage of the two concerned: since, as the Philosopher explains, one needs the goods of the other, and vice versa. Now, what is done for common benefit must not weigh more on one than on the other. Hence, the reciprocal contract must be based on equality. But the value of the things that serve man is measured according to the price that is given: for which, as Aristotle says, it was invented money. [. . .] Second, we can consider the sale as much, accidentally, it constitutes a gain for one and a loss for the other: e.g., when one urgently needs something, and the other is harmed by depriving himself of it. In this case, the right price should not be defined only by looking at what is being sold, but also at the damage that the seller suffers from the sale. And so you can sell for a price higher than the intrinsic value of the thing, even though you don't sell more than it is worth to the owner. And if one receives a significant advantage from the purchase, without the seller being harmed by depriving himself of what he sells, he has no right to increase the price. As the buyer's advantage does not depend on the seller, but on the condition of the buyer: now no one has to sell to another things that do not belong to him, although he can sell the damage he himself suffers. However, those who obtain a significant advantage from the purchase can increase the compensation of their own free will: and it is a sign of nobility of spirit." (ST, II-II, q. 77 a. 1, our own translation from the Italian version reported in **Supplementary Appendix 2**).

In St. Thomas, as for other human activities, also the sale is qualified by its purpose, which in this case is the common advantage. Therefore, the value of things (commodities) must be measured for the advantage they procure (and also for the possible damage that the sale entails to the seller), more than for their labor-value. The Catholic culture of the Middle Ages therefore emphasized the subjective satisfaction to which the economic good must respond, reflecting on the aspect of the final cause of the use of the good itself. This seems a precursor of the Austrian school of economics according to which the value of a good is given by the importance that is *subjectively* attributed to it. Work in the middle age was an element of determining the value of things produced. In specifying the price, quality and quantity were taken into account, and the qualification of the subjectivity of the work itself was also calculated in reference to the social class to which the worker belonged (see Sapori, 1932; Barrera, 1997; Schlag, 2020).

The Just Price, the Subjectivity of Value and the Social Doctrine of the Church

For the interpretation of the Gospel of Matthew according to St. Thomas, the application of the "just price" to the "just wage" for day laborers can help us, since the wage is the price of labor. For this purpose, the classification of the different types of work according to Mises (1949) may be useful. Mises' distinction between "introversive labor" and "extroversive labor" points out to the relevance of the differences existing between workers and the importance of considering their motivations and personal attitude or skills. An interesting development of Mises' *subjective theory of value* can be found in Aranzadi del Cerro (2020). This characterization of work seems consistent with the social doctrine of the Catholic Church (Pontificio Consiglio della Giustizia e della Pace, 2005, p. 151) which in part III "The dignity of work" discusses the subjective and objective dimension of work. No. 270 states that:

> "human work has a double dimension: objective and subjective. In an objective sense it is the set of activities, resources, tools and techniques that man uses to produce, to dominate the earth, according to the words of the Book of Genesis. Work in the subjective sense is the action of man as a dynamic creature, capable of carrying out various actions that belong to the process of work and that correspond to his personal vocation: "man must subjugate the earth, he must dominate it, because as an "image of God" he is a person, that is, a subjective creature capable of acting in a programmed and rational way, capable of deciding about himself and tending to realize himself (Pontificio Consiglio della Giustizia e della Pace, 2005, p. 151, our translation)."

The subjective dimension of work, that is a stable dimension, must have priority over the objective one which is contingent. Subjectivity gives work its peculiar dignity which prevents it from being considered as a mere commodity. Interestingly, No. 273 deals with the "just evaluation" of work reporting the following text, taken from the Lett. Enc. of Pio XI *Quadragesimo Anno* AAS 23 (1931) 200:

> "Work cannot be evaluated with justice if its social nature is not taken into account: "since if there is not a truly social and organic body, if a social and juridical order does not protect the exercise of work, if the various parts, one dependent on the other, are not connected to each other and are not mutually accomplished, if what is more, they do not associate, as if to form a single thing, the intelligence, capital, labor, human activity cannot produce its fruits, and therefore it will not be possible to evaluate it with justice or to remunerate it adequately, where its social and individual nature is not taken into account (Pontificio Consiglio della Giustizia e della Pace, 2005, p. 152, our translation)".

PRELIMINARY RESULTS

Extension of Arguments to Research Evaluation

What do the two parables of the Gospel of Matthew interpreted in the light of the thought of St. Thomas say today to the evaluation of the research? In real life academic settings, the landowner should be *accountable* for the ways he spends his credits (money or career advancement opportunities) and the quantity of available credits is *finite*. The Kingdom of God, salvation, on the other hand, is an *infinite* good and here the landowner gives the *same* salary to different marginal products without budget constraints, by applying a (very peculiar) distributive justice based on the grace that is offered to everyone unconditionally.

In this paper we try to unhinge the idea of merit connected to the well-known Matthew's effect of the sociology of science. Considering merit as the simple reward for productivity, according to the parable of talents, is inappropriate to give bibliometric indicators an objective epistemic value. Bibliometric indicators are infused with meaning through assessment practices in specific contexts and used as "judgment devices" as illustrated by Hammarfelt and Rushforth (2017). In Matthew's Gospel there is also the parable of vineyard workers which is useful to counterbalance the parable of the talents. Through the parable of the vineyard workers and the consideration of the social justice in St. Thomas' vision, our proposal is to make a *more balanced* evaluation than an evaluation that considers only scientific productivity.

Why do we have to apply a *just price*, *control and counterbalance* productivity indicators in performance assessment? Because we recognize, as recalled in the *Introduction*, that there may be stochastic components, luck, related to multiplicative phenomena connected to *publish or perish* inflationary mechanisms at the base of productivity indicators' distributions.

Research is a complex activity which is uncertain. Research is a classic public good which is non-excludable (it is not possible for a user to exclude others from using the good) and non-rivalrous (when one person uses the good he/she does not prevent others from using it).

Here we consider research as a social practice and adopt MacIntyre's definition of social practice that is defined on the basis of peculiar internal goods, i.e. research objectives and the criteria of excellence that concern them, and of the psychological characteristics of the researchers that make them possible (MacIntyre 1985). Gläser and Laudel (2015) proposes an interesting social characterization of researchers' career distinguishing among the cognitive dimension, the scientific/disciplinary communitarian dimension and the organizational/institutional dimension. We consider performing research evaluation as a social practice that should take into account the social dimensions in which researchers and their research practices are embedded in. We need then a *superior model of equity*, a form of social justice to mitigate the asymmetries of bibliometric indicators that can be unfair if used alone.

Returning to the parable of the workers, if the production function is unique, i.e. the same for all, then the first hour workers are right to get angry with the landowner. However, if there are different functions of production, each has its own, the merit and remuneration has to do with subjectivity (see Section *Justice in the Thought of St. Thomas*). We need to know many things about the worker, not just how much he produces. To make justice to the individual, productivity indicators alone are not enough. We may use bibliometric indicators as *minimum thresholds* that must be accompanied by other personal/social characteristics, included e.g. in his\her curriculum vitae (Gläser and Laudel, 2015).

In the parable of workers, a more general notion of merit is applied, which includes a stochastic, random component, defined grace in the philosophical-theological context of St. Thomas. Grace is stochastic because it is offered to everybody but to apply individuals must adhere to it. From this freedom of choice to adhere to it comes the stochastic component of grace.

We can apply by imperfect analogy the logic of the parable of the vineyard workers to research, that we can consider as an infinite good which has a component of *serendipity* (making discoveries by chance and finding an unexpected unsolicited thing while looking for something else, see Merton and Barber, 2004). It is therefore necessary to leave some space in the system, not to reward only the effort because research is a non-standard production activity, which includes stochastic components. Creativity, for example, is favored by effort but also by waste.

This consideration leads us to think that in order to being able to unambiguously interpret and use bibliometric indicators in research evaluation we have to consider also *individual* and *social* characteristics of researchers. Building on the notion of practice of MacIntyre (1985), in Daraio and Vaccari (2020) we argue that the most appropriate level of analysis for building a "good evaluation" is that of "research practice," intended as a form of social practice. The adoption of this level of analysis requires a paradigm shift in the assessment of research from an evaluation centered only on products (outputs of the research, i.e. papers and citations) to an evaluation focused on the functions of research practices, i.e. taking also the process of research/knowledge production into account. Recognizing the importance of the process of production of research has important implications also for the *management of research evaluation*. According to the scheme proposed by Ouchi (1979) for the design of organizational control mechanisms it is necessary to consider two characteristics of the realized activity: i) ability to measure the output and ii) knowledge of the transformation process. Research is characterized by the low ability to measure the output and imperfect knowledge of the transformation process. For this type of activity, the form of organizational control suggested by Ouchi (1979) is that of the *clan,* or network using a more current synonym (not the market or the bureaucratic hierarchy). The social prerequisites of clan control are the most challenging and include "shared values" and "beliefs". The same is true for the organizational control of the people of the clan, which is based on the *identification* of the person with selection/screening and training on both skills and values.

Epistemic Foundations of the (Multiple) Concepts of Merit

The theological reflection on merit, understood as the remuneration due to an action or conduct (Colom and Rodríguez Luño, 2003, p. 219-222), allows interesting insights. Although Sacred Scripture uses the human concept of remuneration to express the reality of merit, as we have seen in the parables of the Gospel of Matthew, the meaning of this biblical notion goes beyond the human idea of reward. On the basis of the content of the merit, theology distinguishes between merit *de condigno*, that is due in justice, and merit *de congruo*, which presupposes a certain convenience, but taking into account the donor's liberality. In this latter merit, which does not arise from a proportionality between acting and the reward, but from the pure liberality of the donor, grace enters.

The application of St. Thomas's theory of the just price of work, developed in the social doctrine of the Catholic Church cited in the previous section, suitably revised in the light of the evaluation context, can help us understand if and under which conditions bibliometric indicators represent the just compensation for the research work carried out. Considering the "subjective" nature of the just price or salary, the inclusion of the personal characteristics, stable traits of character and motivations of researchers, their epistemic *virtues* (considered as the intellectual virtues embodied in the communities of researchers; see Turri et al. (2019) for an overview on the recent philosophical debate on virtue epistemology), certainly play a relevant role. The intrinsic social dimension of work, which is present in the social doctrine of the Catholic Church, highlights that working is increasingly work with others and for others. Even the fruits of work offer an opportunity for exchanges and relationships.

Using the thought of St. Thomas, we can broaden our perspective instead of considering the classical "nature and nurture" considering "nature, grace and nurture" this allows us to include merit in a broader ontological context which include grace and mercy together with justice. By studying Christian theology and drawing on St. Thomas, we can have a broader explanation, that does not contradict our reason, but at the same time transcends what we can grasp with an exclusive use of it. Our thesis is that Christian theology, based on the systematic thought of St. Thomas, can help us to clarify the complexity of the concept of merit. In fact, merit is a concept connected with many others: in the *Summa,* in a theological context, St. Thomas inserts it within grace. In non-theological terms, we can say that merit is also connected to gratitude. It is certainly connected to the concepts of justice and mercy as illustrated above, and with the consideration of other personal aspects. The ontological framework offered by Christian theology is a rich one, suitable to find out and reconcile different concepts of justice within a reasonable, logical and systematic ordered system.

CONCLUDING REMARKS AND FURTHER RESEARCH

In this work we tackle the ambiguity of bibliometric indicators, that is, of what they measure and in particular whether they measure the merits of researchers rather than luck or chance, starting with the clarification of the concept of merit.

Using philosophical argumentation, we attempt to show the usefulness of Christian theology, or the science of faith, to clarify the concept of merit, overcoming its conceptual ambiguities. From the analysis carried out, based on the thought of St. Thomas, the *subjectivity* of the "just evaluation" emerges and this requires the inclusion, in the notion of merit, of the personal characteristics, stable traits of character and motivations of researchers, in other words, the epistemic *virtues* that are generated in the research practices, conceived as social practices. In order to give an unambiguous interpretation to bibliometric indicators, it seems necessary to include and account for these subjective characters or virtues of scholars in the evaluation, definitely broadening the evaluation perspective. Moving from an evaluation based only on the output or results (e.g. counting only number of papers and citations) to an evaluation that considers also the process of production of research, the virtues and motivations of individuals who take part in the social research practices and other qualitative information included for instance in scholars' curricula.

The considerations reported in this paper are still at their infant stage and need further research toward a *systematic conceptualization* of merit. Further research is also needed to understand and explain the connection of merit, its assessment and performance evaluation.

There are many other aspects that remain to be explored further in an attempt to understand who are "good" researchers, what makes "good" a good researcher and how to make a "good" evaluation of researchers. Among these, an interesting track to follow is the philosophical-theological study of the nexus between *effort and luck*, considering the initial conditions (natural talents) and the contextual factors. Deepening the theological knowledge of the relationship between *merit and grace* that we have introduced in this paper could help us to dissect the relationship that exists between *effort and luck* in scientific performance because, as we have seen, grace has a stochastic component similar to the luck that is offered to all but it is necessary to adhere to it. In addition, in theology, according to the logic of God, it is important not only "how much" one does but also "how" one does. According to James (2:26) e.g. "faith without works is dead." The deepening of the knowledge of God's logic through the science of faith hence could give us interesting insights on the relevant relationship between *quantitative* and *qualitative* dimensions of research performance.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

The funding support of Sapienza Awards n. PH11715C8239C105 and n. RM11916B8853C925 is gratefully acknowledged.

ACKNOWLEDGMENTS

This essay has been written within the Joint Diploma St. Thomas Aquinas Doctor Humanitatis 2019–2020. Warm thanks go to Marco Vanzini and Ángel Rodríguez Luño for their valuable discussions and suggestions on the overall paper and in particular on the theological aspects of work. A special thank goes to Andrea Bonaccorsi for his deep and helpful discussions and suggestions on the overall paper and in particular for the extension to research evaluation of the main arguments of the paper. Many thanks also to Daniele Checchi, Giancarlo Ruocco and Alessio Vaccari for their interesting comments and suggestions. Finally, our deepest acknowledgment of the useful and constructive comments and suggestions of two anonymous reviewers that allowed us to greatly improve our work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frma.2020.614016/full#supplementary-material.

REFERENCES

Abelhauser, A., Sauret, M. J., and Gori, R. (2011). *La Folie Évaluation: Les Nouvelles Fabriques De La Servitude*: Fayard/Mille et une nuits. Paris, Librairie Arthème Fayard.

Albarrán, P., Crespo, J. A., Ortuño, I., and Ruiz-Castillo, J. (2011). The skewness of science in 219 sub-fields and a number of aggregates. *Scientometrics*. 88 (2), 385–397. doi:10.1007/s11192-011-0407-9

Allison, P. D. (1980). Inequality and scientific productivity. *Soc. Stud. Sci*. 10 (2), 163–179

Allison, P. D., Long, J. S., and Krauze, T. K. (1982). Cumulative advantage and inequality in science. *Am. Socio. Rev*, 615–625

Allison, P. D., and Stewart, J. A. (1974). Productivity differences among scientists: evidence for accumulative advantage. *Am. Socio. Rev*, 596–606

Aranzadi del Cerro, J. (2020). "Antropologia, economia e teoria dell'azione di Mises," in *Storia del pensiero economico e fede cristiana*. Editors M. Schlag and A. Roncella (Roma, Italy: EDUSC)

K. Arrow, S. Bowles, and S. N. Durlauf (Editors) (2000). *Meritocracy and economic inequality*. Princeton, NJ: Princeton University Press

Barrera, A. (1997). Exchange-value determination: scholastic just price, economic theory, and modern catholic social thought. *Hist. Polit. Econ*. 29 (1), 83–116

Bell, D. (1972). Meritocracy and equality. *Publ. Interest*. 29, 29.

Berg, M., and Seeber, B. K. (2016). *The slow professor: challenging the culture of speed in the academy*. Toronto, Canada: University of Toronto Press.

M. Biagioli and A. Lippman (Editors) (2020). *Gaming the metrics: misconduct and manipulation in academic research*. Cambridge, MA: MIT Press

Brown, C. (2001). Review of Meritocracy and economic inequality edited by kenneth Arrow, sameul Bowles, and steven durlaf. *J. Econ. Lit*. 39 (1), 93–104. doi:10.1257/jel.39.1.93

Carson, J. (2020). Quantification–affordances and limits. *Scholarly Assessment Reports*. 2 (1), 1–10. doi:10.29024/sar.24

Carson, J. (2007). *The Measure of Merit: Talents, Intelligence, and Inequality in the French and American Republics, 1750–1940*. Princeton, NJ: Princeton University Press

Castilla, E. J., and Benard, S. (2010). The paradox of meritocracy in organizations. *Adm. Sci. Q*. 55 (4), 543–676. doi:10.2189/asqu.2010.55.4.543

Colom, E., and Rodríguez Luño, Á. (2003). *Scelti in Cristo Per Essere Santi: Elementi di Teologia Morale Fondamentale. Terza ed*. Roma, RM, Italy: Edizioni Università della Santa Croce

Dahler-Larsen, P. (2014). Constitutive effects of performance indicators: getting beyond unintended consequences. *Publ. Manag. Rev*. 16 (7), 969–986. doi:10.1080/14719037.2013.770058

Daniels, N. (1978). Merit and Meritocracy: *Philosophy & Public Affairs*, 206–223

Daraio, C., and Vaccari, A. (2020). Using normative ethics for building a good evaluation of research practices: towards the assessment of researcher's virtues. *Scientometrics*, 1–23. doi:10.1007/s11192-020-03658-4

De Rijcke, S., Wouters, P. F., Rushforth, A. D., Franssen, T. P., and Hammarfelt, B. (2016). Evaluation practices and effects of indicator user. A literature review. *Res. Eval*. 25, 161–169. doi:10.1093/reseval/rvv038

De Solla Price, D. J. (1963). *Little Science, Big Science and Beyond*. New York, NY: Columbia University Press.

Del Rey, A. (2013). *La Tyrannie de l'évaluation*. Paris, France: La découverte.

D'Aquino, T. (2018). *Commento al Vangelo secondo Matteo Capitoli 15-28, trad. it. 2018, Dall'edizione Latina a Cura di Raffaele Cai o.p. Marietti Torino, 1951*. Bologna, Italy: Edizioni Studio Domenicano

D'Aquino, T. (2014). *La Somma Teologica, tr. it., 2014, 4 voll. Testo latino dell'Edizione Leonina Integralmente Rivisto. Traduzione Italiana: Curata da Tito Sante Centi, Roberto Coggi, Giuseppe Barzaghi, Giorgio Carbone*. Bologna, Italy: Edizioni Studio Domenicano.

Festinger, L. (1954). A theory of social comparison processes. *Hum. Relat*. 7, 117–140

Frank, R. H. (2016). *Success and Luck: Good Fortune and the Myth of Meritocracy*. Princeton, NJ: Princeton University Press

Gingras, Y. (2016). *Bibliometrics and Research Evaluation: Uses and Abuses*. Cambridge, MA: MIT Press

Gläser, J., and Laudel, G. (2015). *The Three Careers of an Academic*. Berlin, Germany: TU Berlin, Center for Technology and Society.

Hammarfelt, B., and Rushforth, A. D. (2017). Indicators as judgment devices: an empirical study of citizen bibliometrics in research evaluation. *Res. Eval*. 26 (3), 169–180. doi:10.1093/reseval/rvx018

Heneman, R. L. (2002). *Merit Pay. Strategic Reward Management: Design, Implementation, and Evaluation*. Connecticut: nformation Age Publishing, 379–393

Jones, T. M. (1994). Two cheers for Meritocracy. *Proc. Int. Assoc. Bus. Soc*. 5, 49–59

S. B. Kaufman (Editors) (2013). *The complexity of greatness: Beyond talent or practice*. Oxford, United Kingdom: Oxford University Press).

Littler, J. (2018). *Against meritocracy: culture, power and myths of mobility*. Abingdon-on-Thames, UK: Taylor & Francis.

Lotka, A. J. (1926). The frequency distribution of scientific productivity. *J. Wash. Acad. Sci*. 16 (12), 317–323

MacIntyre, A. (1985). *After Virtue*. London, United Kingdom: Duckworth.

Maggioni, B. (2009). *Il racconto di Matteo*. Assisi, Italy: Cittadella Editrice.

McNamee, S. T., and Miller, R. K., Jr (2014). *The Meritocracy Myth*. Plymouth. Washington: Rowman & Littlefield Publishers.

Merton, R. K., and Barber, E. (2004). *The Travels and Adventures of Serendipity*. Princeton, NJ: Princeton University Press.

Merton, R. K. (1973). *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago, IL: University of Chicago press

Mijs, J. J. B. (2019). The paradox of inequality: income inequality and belief in meritocracy go hand in hand. *Socio-Economic Review, mwy051*. doi:10.31235/osf.io/dcr9b

Miller, D. (1996). Two cheers for meritocracy, *J. Polit. Philos*. 4 (4), 277–301.

Mises, L. V. (1949). *Revised ed. 1963, Human Action: A Treatise on Economics*. Rome, Italy: Yale University (Italian translation of 2016, L'azione Umana, Trattato di Economia, Rubbettino Editore.

Mondin, B. (2000). *Dizionario Enciclopedico del Pensiero di San Tommaso d'Aquino*. Bologna, Italy: Edizioni Studio Domenicano.

Muller, J. Z. (2018). *The Tyranny Of Metrics*. Princeton, NJ: Princeton University Press.

Nielsen, S. B., and Hunter, D. E. (Editors) (2013). *Performance management and evaluation: new directions for evaluation*. Hoboken, NJ: John Wiley & Sons Vol. 114.

Ouchi, W. G. (1979). A conceptual framework for the design of organizational control mechanisms. *Manag. Sci*. 25 (9), 833–848

Pontificio Consiglio della Giustizia e della Pace (2005). *Compendio della Dottrina Sociale della Chiesa*. Città del Vaticano: Libreria Editrice Vaticana

Porro, P. (2012). *Tommaso d'Aquino un Profilo Storico–Filosofico*. Roma, RM: Carocci

Ruiz-Castillo, J., and Costas, R. (2014). The skewness of scientific productivity. *J. informetrics*. 8 (4), 917–934. doi:10.1016/j.joi.2014.09.006

Ruocco, G., Daraio, C., Folli, V., and Leonetti, M. (2017). Bibliometric indicators: the origin of their log-normal distribution and why they are not a reliable proxy for an individual scholar's talent. *Palgrave Communications*. 3 (1), 1–8. doi:10.1057/palcomms.2017.64

Sandel, M. J. (2020). *The Tyranny of Merit: What's Become of the Common Good?*. London, United Kingdom: Penguin Books

Sapori, A. (1932). Il giusto prezzo nella dottrina di san Tommaso e nella pratica del suo tempo. *Arch. Stor. Ital*. 90 (3), 3–56

Schlag, M. (2020). "L'etica economica di Tommaso D'Aquino," in *Storia del pensiero economico e fede cristiana*. Editors M. Schlag and A. Roncella (Roma, Italy: EDUSC)

Seglen, P. O. (1992). The skewness of science. *J. Am. Soc. Inf. Sci*. 43 (9), 628

Sen, A. (2000). "Merit and justice," in *Meritocracy and economic inequality*. Editors K. Arrow, S. Bowles, and S. Durlauf (Princeton, NJ: Princeton University Press), 5–16

Simon, R. L. (1974). Equality, merit, and the determination of our gifts. *Soc. Res*, 492–514

Son Hing, L. S., Bobocel, D. R., Zanna, M. P., Garcia, D. M., Gee, S. S., and Orazietti, K. (2011). The merit of meritocracy. *J. Pers. Soc. Psychol*. 101 (3), 433. doi:10.1037/a0024618

R. J. Sternberg and S. B. Kaufman (Editors) (2011). *The Cambridge handbook of intelligence*. Cambridge, United Kingdom: Cambridge University Press

Turri, J., Alfano, M., and Greco, J. (2019). "Virtue epistemology," in *The stanford encyclopedia of philosophy*. Editor E. N. Zalta. Available at: https://plato.stanford.edu/archives/fall2019/entries/epistemology-virtue/ (Accessed November 25, 2020).

Vanzini, M. (2019). *La Teologia, Sapere della Fede e Il Suo Carattere Scientifico, Joint Diploma S. Tommaso D'Aquino*. Roma, Italy: Pontifical University of the Holy Cross

Vidaillet, B. (2013). *Évaluez-moi!. Évaluation au travail: les ressorts d'une fascination*. Paris, France: Le Seuil

Xie, Y. (2014). Inequalities in science. *Science*. 344 (6186), 809–810. doi:10.1126/science.1252743

Young, M. D. (1958). *The Rise of the Meritocracy*. Piscataway, NJ: Transaction Publisher

Young, M. (1994). Meritocracy revisited. *Society*. 31 (6), 87–89

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Research Metrics and Analytics

ORIGINAL RESEARCH
published: 09 February 2021
doi: 10.3389/frma.2020.628703

Check for updates

Interpreting Bibliometric Data

Martin Szomszor[1†], *Jonathan Adams*[1*†], *Ryan Fry*[2†], *Chris Gebert*[2†], *David A. Pendlebury*[2†], *Ross W. K. Potter*[1†] *and Gordon Rogers*[1†]

[1]Institute for Scientific Information, Clarivate, London, United Kingdom, [2]Institute for Scientific Information, Clarivate, Philadelphia, PA, United States

OPEN ACCESS

Edited by:
Gernot Deinzer,
University of Regensburg, Germany

Reviewed by:
Domingo Docampo,
University of Vigo, Spain
Carlos Olmeda-Gómez,
Universidad Carlos III de Madrid, Spain

Correspondence:
Jonathan Adams
jonathan.adams@clarivate.com

†ORCID:
Martin Szomszor
https://orcid.org/0000-0003-0347-3527
Jonathan Adams
https://orcid.org/0000-0002-0325-4431
Ryan Fry
https://orcid.org/0000-0002-2388-7359
Chris Gebert
https://orcid.org/0000-0003-1037-9985
David A. Pendlebury
https://orcid.org/0000-0001-5074-1593
Ross W. K. Potter
https://orcid.org/0000-0002-1338-5910
Gordon Rogers
https://orcid.org/0000-0002-9971-2731

Specialty section:
This article was submitted to Research Assessment, a section of the journal Frontiers in Research Metrics and Analytics

Received: *13 November 2020*
Accepted: *17 December 2020*
Published: *09 February 2021*

Citation:
Szomszor M, Adams J, Fry R, Gebert C, Pendlebury DA, Potter RWK and Rogers G (2021) Interpreting Bibliometric Data. Front. Res. Metr. Anal. 5:628703. doi: 10.3389/frma.2020.628703

Many academic analyses of good practice in the use of bibliometric data address only technical aspects and fail to account for and appreciate user requirements, expectations, and actual practice. Bibliometric indicators are rarely the only evidence put before any user group. In the present state of knowledge, it is more important to consider how quantitative evaluation can be made simple, transparent, and readily understood than it is to focus unduly on precision, accuracy, or scholarly notions of purity. We discuss how the interpretation of 'performance' from a presentation using accurate but summary bibliometrics can change when iterative deconstruction and visualization of the same dataset is applied. From the perspective of a research manager with limited resources, investment decisions can easily go awry at governmental, funding program, and institutional levels. By exploring select real-life data samples we also show how the specific composition of each dataset can influence interpretive outcomes.

Keywords: bibliometrics, responsible metrics, data interpretation, research assessment, research policy

INTRODUCTION

In this paper, and in the context of good and responsible research evaluation, we review the challenge of making correct use and interpretation of the rich information on research activities and outcomes that can be mined from the data around academic journal publications and their citations. This challenge exists at three levels. First, summary citation metrics are usually insufficient to enable fully informed interpretation by the intended users, who are typically research experts in their own fields but unfamiliar with the nature of these data. Second, bibliometric analysis is a tool, the good use of which depends on the user and on the context, and it sharpens questions more often than provides answers (Moed, 2020). Third, because alternative visualisations supporting better interpretation require additional work by these users, they often default to simpler metrics because of time pressure.

We consequently suggest that the priority around scientometric research and practice is not about academic development, which has been extensive over the last few decades, but about practical user focus. There is a need for a structured change in describing how best to use bibliometric analysis. The user needs to be able to start from the context of application with an evaluation framework within which they can specify the data and analytical tools relevant to the questions they pose. The user needs to access information that enables an *a priori* understanding of how they will use these tools, so bibliometric researchers must understand and consider the user perspective. And the user needs to see data presented not as summary point metrics but in a form that allows accessible, interpretive exploration. We examine these challenges through analyses of international research activity and performance.

ORIGINS

It is widely acknowledged that bibliometric indicators have become one of the most frequent tools of normal practice in evaluative research management. The development of research evaluation practices has been well documented as it shifted from structured processes drawing on strong peer review (Gibbons and Georghiou, 1987) through strategic reorientation (Georghiou, 1995) to systems increasingly drawing on data and metrics (Martin, 1996; Adams et al., 2007; Hicks, 2010; Hicks and Melkers, 2012. Bibliometric indicators, which had been of a specialist nature prior to 1990 (e.g., Narin, 1976; Martin and Irvine, 1983), were introduced to a wider audience during the 1990s when the former Science Citation Index became accessible on-line as the *Web of Science*™ (WoS). Data used for national comparisons of research performance were published in widely-read journals (May, 1997; Adams, 1998) that brought them to the attention of a larger audience who applied them to institutional, program and policy purposes.

Research evaluation may be seen as a reflection of a broader societal shift to institutional managerialism and public sector accountability. As Langfeldt et al. (2020) note: "mechanisms for constituting research quality notions that were once reserved for highly professionalized knowledge communities have extended to encompass notions generated within policy and funding domains." It was spawned by a growth in research and tertiary education systems that was more rapid than the growth of resources that governments were willing to allocate. For the United Kingdom, as an example with which we have particular familiarity, most projects submitted to Research Councils in the 1970s and peer reviewed as 'alpha - fundable' could be financially supported. Then the proportion funded began to fall, so new but still peer-selective criteria were introduced (alpha 1–5). At the same time the country faced an economic and energy crisis, so wider resource constraints appeared. The bodies responsible for funding research in United Kingdom universities (general grants via the University Grants Committee, project grants via the Advisory Board for the Research Councils) reported on the need for selectivity in research distribution (UGC, 1982; ABRC, 1983; UGC, 1984). Thus, the first national Research Selectivity Exercise was introduced in 1986 and led to a more formalized and structured Research Assessment Exercise (RAE) from 1992.

Such an exercise as the RAE had a profound effect on the strategic view of the research enterprise, the management of research in universities, and the spotlight thrown on the individual researcher. The United Kingdom's procedures also attracted widespread international attention, if not always emulation. It also produced a formidable workload for assessment panel members, who had other full-time roles in addition to the peer review work. Analysis of the results of successive RAEs in 1992 and 1996 were soon augmented with the more accessible bibliometric data then available and thus attention inevitably turned to the idea that quantitative analysis might substitute for some of the onerous qualitative review. After RAE2001, the proposal for a 'metrics based' review process was brought under serious central review (Roberts, 2003) but rejected after a pilot exercise prior to RAE2008 (Evidence, 2009).

The United Kingdom's experience of assessment and metrics' policy was reflected elsewhere in Europe, notably in the Netherlands and Scandinavia and, in due course, the ideas spread (Sivertsen, 2017). As a consequence, research evaluation using, to a very variable degree, some form of publication and/or citation data is now widespread and present in different forms and at various levels in for example: European programs (European Science Foundation, 2012), in Australia (ARC, 2019), Finland (Lahtinen et al., 2005), Italy (CIVR, 2006; Abramo and D'Angelo, 2015), New Zealand (Buckle and Creedy, 2019; PBRF, 2020), Sweden (Karolinska Institute, 2010), Spain (Jiménez-Contreras et al., 2003), Norway (Sivertsen, 2018), the United Kingdom (REF, 2020) and the United States (National Institutes of Health, 2008). Thomas et al. (2020) recently reviewed 350 research papers on performance-based research evaluation arrangements and discuss important limitations in applying and using such research.

Problems

Jappe et al. (2018) noted that there is a gap between the demand for indicator-based performance assessment by research organizations and the researchers being assessed. Researchers - and their works - come from a multi-modality of disciplines and cultures with their own norms and expectations. However, because the academic sector, at discipline level, has taken little or no responsibility for understanding and interpreting quantitative indicators based on citation data, *de-facto* and generic standards of research excellence have been defined at system level by others (including scientometricians and data providers) without being challenged by the implied authority of the domain experts. While the possible forms of analysis are diverse, Jappe (2020) reviewed 138 evaluation studies from 21 EU countries, covering the period 2005 to 2019, and found that bibliometric research assessment, which was common to the United Kingdom, the Nordic network, the Netherlands and Italy, was most frequently based on 'citation impact' metrics, usually with reference to international scientific fields.

The most widely used standard indicator for 'citation impact' is the number of citations received by a publication, normalized "with reference to international scientific fields" (Jappe, 2020). It is generally understood that papers with higher citation counts are associated with greater influence or 'impact' since they reflect acknowledgment by other researchers (Garfield, 1955). Citation counts have in turn been shown to be correlated with other indicators of research performance, such as peer review (Evidence, 2007; Waltman, 2016; Aksnes et al., 2019).

To this simple summary several essential caveats must be made. First, the citation metric is only an indicator of impact. Citation counts reflect impact indirectly through a general population relationship and, for individual papers, may be awry in their information. Indeed, the mantra 'on average' has wide applicability to every aspect of this kind of analysis. Second, citation counts rise over time, older papers have more citations 'on average' than more recent and an adjustment must be made to take account of the years since publication. Third, citations accumulate at rates that are field dependent. For example, life sciences are more prolific and exhibit higher rates of citation on average than technological and social sciences and an adjustment must be made to take account of the field of publication (Moed et al., 1985a). Fourth, document type affects citation rates with reviews in journals cited more often than articles ('on average', see Ketcham and Crawford, 2007; Miranda and Garcia-Carpintero, 2018) while conference proceedings are cited less often than journal papers.

The 'standard indicator' (the observed document citation count) is therefore processed before analysis. It is, usually, compared to the global average (or 'expected') count for the same document type, year of publication and field. Field is usually determined from a pre-set categorical structure which, for WoS, is based on journal assignment to discipline-based categories. Then, the ratio of observed/expected citation counts is used to calculate an average Category Normalized Citation Impact (CNCI) for a research group, institution or country. Again, recall that this CNCI value is an indicator, not a metric, and is now at some distance from the target research activity under evaluation.

So, this general procedure refers to a simple index, inferred to be a reasonable indicator of other aspects of research performance for larger samples (Rogers et al., 2020), that may or may not be relevant to the research objectives that are the proper target of an evaluation. For the humanities, citation counts are of little informational value and indeed journal articles are usually secondary to monographs as a signal of intellectual significance. For applied research of industrial or policy significance, value is reflected in utility and application, not in later academic references. Even where citations are a more appropriate currency, the basic caveats recognized long ago (Moed et al., 1985b), along with a large number of more nuanced issues of qualification (Pendlebury, 2009), are not universally understood by the domain-expert users and their research managers. This leads to extensive misuse (Moher et al., 2018) and consequent reaction from researchers and observers (DORA, 2012; Hicks et al., 2015; Wilsdon et al., 2015).

Users and Criteria

What do research panels and committees do and how do they use (and possibly abuse) bibliometric data and analysis? There is, as Jappe et al. (2018) noted, a gap between these context-specific users and the people who typically explore, analyze and propose the metrics (scientometricians).

One of us (JA) has experience of committee work at national level (as a science policy adviser in the United Kingdom and Australia) and institutional level (as Director of Research Strategy at the University of Leeds), as well as through commercial consultancy with universities in other countries. The key common learning point from these diverse experiences is that research metrics are hardly ever an arbiter in normal practice; they are more typically one of several adjunct sources of information. The information in front of a decision-making group is there to help it to arrive more confidently and speedily at that decision so as to support research management and enable activity to proceed. The presentation of a table of simplistic and opaque metrics is unlikely to do this and it competes for attention with other considerations such as apparent opportunity, real resource constraints, dominant voices, and local and third-party politics.

The United Kingdom's Advisory Board for the Research Councils criteria for scientific priorities (ABRC, 1987) were published as a guide for both Research Council peer reviewers and committees, as well as a general aid to research planning. They draw implicitly on the ideas of Weinberg (1963) and set out criteria, both internal and external for any research project, that have stood the test of time (ABRC, 1987).

A. Internal: i) timeliness - expectation of rapid scientific advance (in 5,10 or 20 years); ii) pervasiveness - likelihood of a wide range of links with other research; iii) excellence.
B. External: i) exploitability - potential for nationally profitable industrial or commercial use (in 5, 10 or 20 years); ii) applicability - potential for uses leading to other benefits: social, environmental or related to Government policy (in 5, 10 or 20 years); iii) significance for education and training.

The ABRC noted that in all judgements, whether internal or external considerations are to the fore, the question of affordability comes into play: the likely benefits of research programmes (as for any other form of public expenditure) must always be weighed against their cost.

Excellence is one among six ABRC criteria and the only one where bibliometric data appear likely to support decision-making more effectively (see Bornmann, 2014). We will show later in this paper that bibliometrics can in fact also throw light on timeliness and pervasiveness. Moed (2005, page 57) also makes the point that citations discriminate best between good and bad but less well between good and excellent. Context, reflected here in the external criteria, is always an essential part of evaluation and Nature (2018) drew attention to the truism that "Excellence depends on context." What is excellent in advancing basic knowledge may not address immediate problems, and vice versa.

These criteria provide a balance of reference points for a working framework (sensu Moed, 2020), which is a fundamental requirement for evaluation. Defining context and purpose provides a framework, or scenario, in which bibliometric analysis is introduced as a purposive tool, almost certainly to improve broader interpretation and understanding, increase confidence in the overall information pool through challenging heuristic assumptions (Bornmann and Marewski, 2019) and thus inspire greater and more rapid progress toward a decision.

A structure for consideration of the context for 'good research' has been proposed by Langfeldt et al. (2020) and they discuss three perspectives from which differences of opinion may arise: 1) research fields vs. policy spaces; 2) 'attributes' of originality/novelty, plausibility/reliability, and value or utility; and 3) 'sites' where quality notions emerge: researchers, communities, organizations, funders and national policy. We agree that it would be valuable to consider how any research project or program would be seen in these perspectives before deciding how best to evaluate the work.

Bibliometric analysis without a clear locus in a contextual and evaluation framework is unlikely to be used effectively. A table of point metrics, for example, has little contextual value since it is unconnected to other aspects of the activity under review. We need instead to move to more complex perspectives, based on multiple points of reference, that explain the purpose, and hence the purposive structure, of the evaluation and enable informed interpretation and comprehension of meaning.

An Example

To illustrate the problem of interpretation that comes from inappropriately simplistic bibliometric information, we start from a table of point metrics, consider what these would show us and then move to other analyses that may reveal alternative or nuanced interpretations. We start with bibliometric indicators for a cross-section of ten countries. Five of these might be considered to have both large and well-funded research economies (United States, China, United Kingdom, Germany and Australia) and the other five, while improving, presently have both relatively weaker funding and smaller research output (**Table 1**).

Data summaries similar to that in **Table 1** can be found in many reports from government agencies and in news media. It will be immediately obvious that it tells us nothing about the subject spread of research, which would be important for any informative analysis, nor about the context of relative research expenditure, human capacity and industrial R&D of any of these countries.

More significantly, from the perspective of the present paper, we see results that are at least likely to raise eyebrows if not actually to induce skepticism about the data source. Does Sri Lanka really have an average CNCI equal to the United States when the latter produces more than 500 times as many publications? What does it mean if Iran has the highest rate of cited papers when it is the second lowest in average CNCI? How, in other words, are these point metrics compiled and calculated?

We can also question the representative nature of 'average' or total values of activity across the period. Annual trends in CNCI for the large, well funded research economies appear to be fairly steady across the decade, improving in three cases albeit drawing attention to a gradual decline for the United States. China has a steady upward trend in impact, and Bulgaria also improves throughout though its smaller output means that its line is more variable. Sri Lanka dives, however, from an exceptional CNCI in 2015 and Indonesia falls from slightly above world average to barely 0.5 of that benchmark in 2019 (**Figure 1**). Evidently factors other than the innate research competence of the economy are at work in these instances of indicator volatility. These both are small research economies, relatively low in their research investment and–as we shall see–highly engaged in international research collaboration.

What do the numbers tell us? The data suggest that average CNCI for at least two of the ten nations is unreliable, since doubts about the relative average for Sri Lanka in **Table 1** seem confirmed by its volatility. Does that introduce doubt about the more stable values? It certainly raises questions about the detail in the evident mass of publications (**Table 1**) that feed the indicators for the larger economies. How representative can a single indicator be when it is chosen to stand for millions of publications and tens of millions of citations? More information is required to properly interpret either a table or a graph of summary metrics. Relevant factors explored over the last 2 decades include data granularity, collaboration, geography, history, national research culture, and accessible visualization of underlying distributions that reveal the broader context of the research under evaluation.

REINTERPRETATION

Granularity and Categorization

The CNCI values shown in **Table 1** and tracked in **Figure 1** are calculated at the level of the journal-based categories used in the *Web of Science* (WoS) of which the 254 current categories cover all subject domains in the sciences and arts. Separately, Clarivate also has an Essential Science Indicators (ESI) classification with 22 broad categories that do not include arts and humanities. The Clarivate InCites platform offers additional options to users, including the popular Australia New Zealand Standard Research Classification system (https://www.arc.gov.au/grants/grant-application/classification-codes-rfcd-seo-and-anzsic-codes) which is a hierarchy of Fields of Research (FoRs) with 22 FoRs at the highest level and then nested fields at increasing granularity, thus: Division 03 Chemical Sciences; Group 0302 Inorganic chemistry; Field 030206 Solid state chemistry. InCites has other classifications such as those used in Brazil by CAPES (Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior) and FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) and some developed for particular purposes, such as the RAE/REF Units of Assessment (UoAs) used in the United Kingdom. All of these have validity and utility in their relevant context, none are either right or wrong, but it is important that users understand which classification they have applied and what its purpose and properties may be.

For example, in calculating CNCI, the citation count for a specific publication is compared to (i.e. normalized against) the world average for the year of publication of all documents of the same type (such as article, review or conference proceeding) in the same Web of Science category as the journal in which the publication appeared. Zitt et al. (2005) drew attention to the possibility that CNCI would change according to the level (described as the 'zoom') at which any normalization occurs. The possible effects of changing the reference point at which normalization is made had also been noted by Hirst (1978) in relation to 'Discipline Impact Factors'; methods for comparing bibliometric indicators across fields have been reviewed by Schubert and Braun (1993, 1996); and Glanzel and Moed (2002) commented on the effect of different levels of aggregation.

To explore how the categorization of the data might influence the type of metrics in **Table 1**, we tested the effect of the 'Zitt zoom' on our perspective of research performance by analyzing the relative impact of articles submitted for assessment in the United Kingdom RAE2001. We compared impact at three different levels of normalization for university departments at the three highest grades (4, 5 and 5*) awarded in three Units of Assessment (UoA13 Psychology, UoA14 Biological Sciences and UoA19 Physics). The outcome was a significant positive correlation between peer judgements and citation impact at some, but not all, levels of data aggregation.

TABLE 1 | Summary metrics for the research production (numbers of documents indexed in the *Web of Science*) and performance (category normalized citation impact, CNCI world average = 1.0) of a global spread of ten countries during a recent ten-year period (2010–2019). Countries are ranked on CNCI.

	Web of Science documents	Average CNCI	Times cited	% Docs cited
United Kingdom	1,981,903	1.41	26,932,154	65.6
Australia	888,127	1.41	12,626,406	72.4
United States	6,838,175	1.31	90,031,964	63.9
Sri Lanka	13,068	1.31	170,284	63.6
Germany	1,615,968	1.30	23,029,125	71.1
Bulgaria	38,366	1.01	360,385	60.2
China	3,743,888	0.99	39,306,476	71.5
Argentina	121,077	0.96	1,321,844	71.4
Iran	362,748	0.91	3,428,680	77.9
Indonesia	85,885	0.81	342,576	39.1

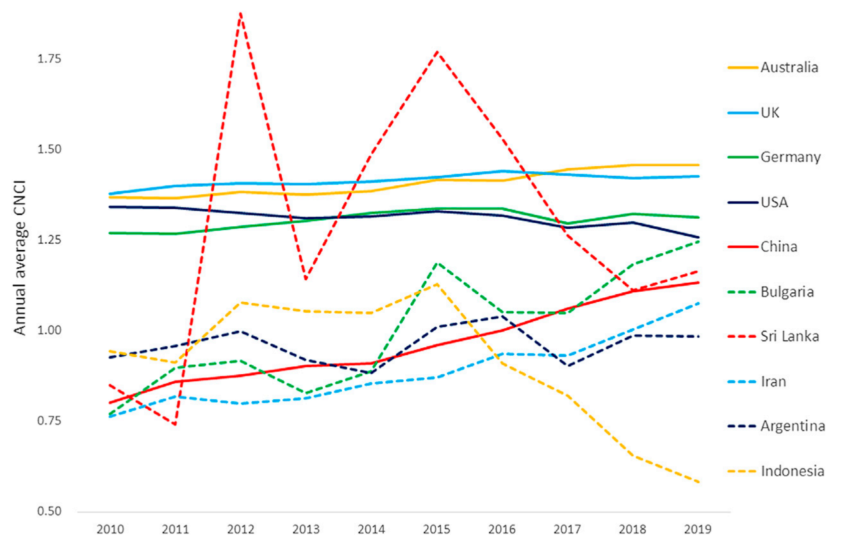

Figure 1 | Annual trends over the last decade in national Category Normalized Citation Impact (CNCI) for the ten countries summarized in **Table 1**.

The citation count for each paper was individually normalized against the average counts–taking note of publication year - for the journal in which it was published, for the WoS category to which the journal was allocated and for the complete data pool for the relevant UoA. When citation counts were normalized at journal level there was little evident difference between performance at any grade, so no link could be made between peer review outcomes and a citation index. But when the normalization was relative to the WoS category or the entire UoA, then on average the higher graded units had a statistically significant higher relative impact. These data support Zitt et al.'s (2005) analysis (**Table 2**).

This has practical implications for research evaluation. The implication is that the material submitted by units that peer reviewers graded at 4 is actually sourced from journals of lower average impact than the material submitted by the units graded at 5 and 5*. Thus, when the level of analysis is relative to journal these items appear to be of similar impact relative to the medium in which they are published. When the viewpoint is zoomed out to the WoS categorical level then the higher absolute citation count for the articles produced by the more highly graded units becomes apparent, and even more apparent at the UoA-level.

The possibility that the level of 'zoom' will affect our assessment of relative impact is an important insight. A clear risk is that very fine-grained assessment becomes self-referential. Clearly, the existence of more than one view and hence more than one interpretation of performance would need to be taken into account in any evaluation methodology. Ideally, the appropriate level of 'zoom' would be independently considered, explored and reported before confidence in the outcome of assessment could be validated. This is likely to be a serious challenge unless a reference indicator is available and will generally require any evaluation to be carried out at multiple levels for a reflective review.

It should also be noted that not all classification systems draw on all available data. The ANZ Fields of Research (FoRs), for example, are used in the 'Excellence in Research for Australia (ERA)' evaluation process where submissions made by universities are assigned to FoRs by reference to expert-assigned journal lists. This results in a marked reduction in the volume of articles and reviews compared with the numbers indexed for any country or institution within the Web of Science. **Table 3** shows the ratio between the total available publication dataset and the number actually assigned

TABLE 2 | The average Category Normalized Citation Impact (CNCI) of articles and reviews published during 1996–2000 by research staff at United Kingdom universities for units graded 4, 5 or 5* in the Research Assessment Exercise 2001 (RAE2001). Data are shown for three Units of Assessment (UoA) with the numbers of units at each grade and the CNCI for their publications with citation counts normalized at three levels of granularity: the journal of publication; the *Web of Science* (WoS) journal category; and the data set for the entire UoA (Adams et al., 2008).

Grade at RAE2001	UoA13 psychology				UoA14 biological sciences				UoA19 physics			
		Average CNCI				Average CNCI				Average CNCI		
	Number of units	Journal based	WoS based	UoA based	Number of units	Journal based	WoS based	UoA based	Number of units	Journal based	WoS based	UoA based
Grade 4	17	1.22	1.40	0.80	17	1.29	2.35	1.89	15	1.28	1.84	1.98
Grade 5	17	1.18	1.80	1.05	30	1.11	2.33	2.33	23	1.47	2.51	2.96
Grade 5*	12	1.32	2.38	1.63	11	1.18	2.53	2.93	5	1.82	3.32	3.75

TABLE 3 | The ratio between numbers of papers assigned to the ten countries listed in **Table 1** via the Web of Science journal-based disciplinary category scheme and six other categorical schema used in Clarivate InCites (schema identified in Note). The variations in the proportion of the literature that is covered will affect both the numerator and denominator citation counts in any subsequent normalization calculation of citation impact (see **Figure 3**).

	ESI	For L1	For L2	REF2014	CAPES49	FAPESP
United States	0.85	0.80	0.64	1.00	1.00	1.00
China	0.78	0.71	0.54	1.00	1.00	1.00
United Kingdom	0.81	0.80	0.65	1.00	1.00	1.00
Germany	0.84	0.77	0.61	1.00	1.00	1.00
Australia	0.86	0.84	0.67	1.00	1.00	1.00
Iran	0.90	0.80	0.64	1.00	1.00	1.00
Argentina	0.90	0.82	0.66	1.00	1.00	1.00
Indonesia	0.34	0.35	0.26	0.98	1.00	1.00
Bulgaria	0.75	0.60	0.48	0.98	1.00	1.00
Sri Lanka	0.76	0.72	0.59	0.99	1.00	1.00

Note: *(ESI = 22 Essential Science Indicators journal categories excluding Arts and Humanities; FOR = ANZSRC Fields of Research where L1 = journals mapped to 24 broad categories and L2 is 212 specific categories nested within L1; REF2014 = 35 of 36 United Kingdom subject panels for Research Assessment Exercise 2014; CAPES = a Brazil schema of 49 evaluation areas used by Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior; FAPESP = 72 categories used by Fundação de Amparo à Pesquisa do Estado de São Paulo, Brazil; PL19 = the Polish schema of 44 categories used for a 2019 evaluation exercise).*

to each country via six other schema. Some schema, especially the journal lists for the ANZSRC Fields of Research, reduce the available data for countries such as Indonesia by as much as half. Even for the United States and the United Kingdom the publication set is down by 20% (the broad L1 categories) or 35% (the specific L2 categories). By contrast, the schema for the United Kingdom's REF and those used in Brazil by CAPES and FAPESP essentially draw on the full source material.

Each scheme has been designed with a particular purpose in mind and draws on and organizes the literature accordingly. The variation in dataset coverage is an intentional outcome of this. However, should the unwary employ a scheme that 'looks right' without recognizing its characteristics then they will obtain a result that may differ from their expectations (**Table 3**).

Categorical schema also have an effect on CNCI, as seen in the 'Zitt zoom' example in **Table 2**. Unsurprisingly, indeed reassuringly, there is a very high degree of correlation between the CNCI values obtained from citation counts normalized under different categorical systems. However, the correlation is not perfect and there can be differences both in the y-intercept, which would move all values up or down, and the slope, which would differentially affect organizations with lower and higher average impact. Matching data categorization to the objectives of the assessment is therefore essential if equity is to be maintained across all parties under assessment.

The average CNCI for all United Kingdom universities (2015–19), taken across all discipline categories in each of several different categorical systems, is shown in **Figure 2**. The effect of moving from WoS journal categories to the FOR 2-digit Level 1 is to depress most institutional CNCIs but this effect is most marked below world average CNCI and almost negligible at the upper end of the distribution. There are also some evident outliers, so the effect is far from uniform. There is a much closer correlation between the CNCI values for the WoS categories and the topical categories created by a citation-based clustering developed by the Center for Science and Technology Studies (CWTS, University of Leiden). Specifically, we used the 'meso' level in CWTS's three tier system. Comparison between the CNCI outcomes using CWTS meso categories and the FOR1 categories shows again that the FOR system depresses the CNCI values. A shift to a finer-grained level, using the CWTS micro and the ANZ FOR Level 2 categories, produces a similar effect but the change in slope is more evident and the depression in the low CNCI part of the distribution is relatively greater (**Figure 2**).

The changes in relative positions for the four tracked universities illustrates the considerable residual variance in these example graphs. The shift from one categorical system to another is never uniform across all the entities. Comparing WoS with FOR1 (**Figure 2A**), there are six universities with an average CNCI of 1.7 when using WoS journal categories that would achieve CNCI values ranging between 1.45 and 1.85 if FOR1 categories were used for data grouping and normalization. Looking at the four tracked universities in comparisons between CWTS-MESO and FOR1 (**Figure 2B**) and between CWTS-MICRO and FOR2 (**Figure 2D**), the highest performer university gains in the shift to FOR but the other three all suffer a reduced CNCI.

These shifts may be due to subject mix, because each system assigns journals differently across the specific category series so the content of global baselines changes, or it may be another, less

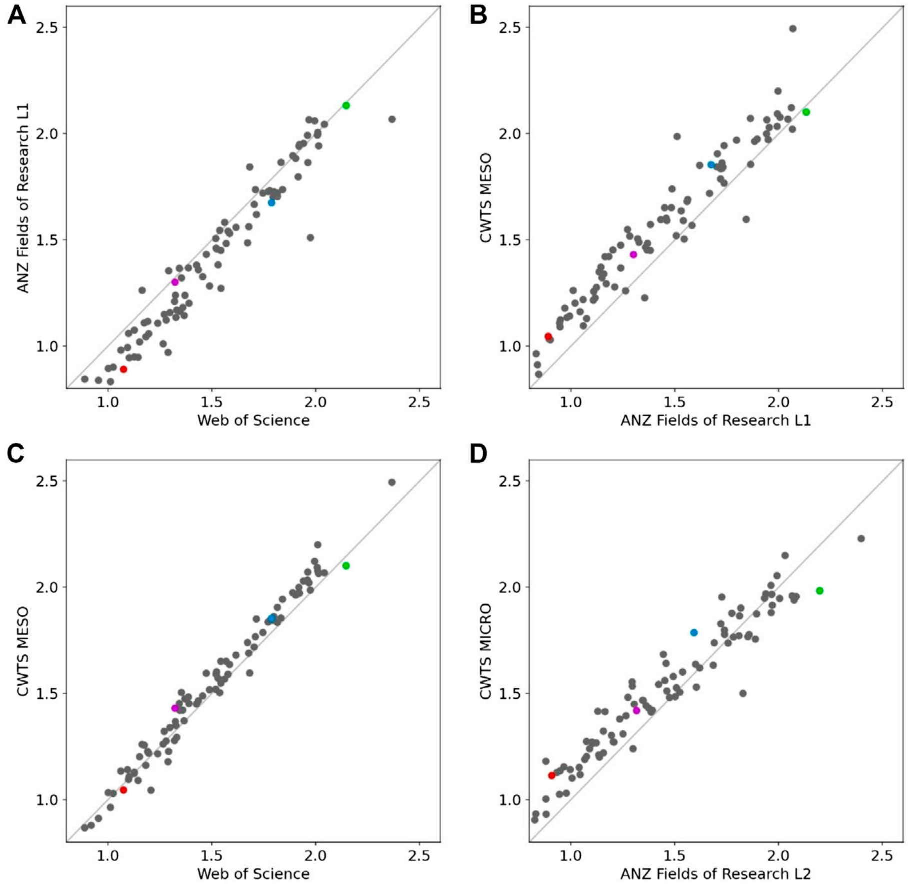

Figure 2 | (A)–(D) Correlations between the average Category Normalized Citation Impact (CNCI) of United Kingdom universities (*Web of Science* indexed publications for 2015–2019) when different schema (see Note) are used to categorize the institutional and global publication data. In each case the correlation is highly significant but the variance about the regression differs for specific institutions. Four universities with distinct research histories and portfolios are highlighted with a constant color point. (*Web of Science* (WoS) categories map journals to 254 fields; ANZSRC Fields of Research (FOR) use L1 = journals mapped to 24 broad categories and L2 = 212 specific categories nested within L1; CWTS MESO and MICRO refer to coarse and fine citation-based categories developed by CWTS, Univ of Leiden). **(A)** FOR1 vs. WoS, $n = 86$, correlation = 0.968. **(B)** CWTS Meso vs. FOR1, $n = 86$, correlation = 0.954. **(C)** CWTS Meso vs. WoS, $n = 86$, correlation = 0.986 **(D)** CWTS Micro vs. FOR2, $n = 86$, correlation = 0.926.

apparent factor, but it materially affects the relative institutional outcomes and cannot be ignored.

The effect of this on the ten countries in **Table 1** reflects these trends and is, in some instances, noticeable (**Figure 3**). The data in **Table 1** (based on WoS journal categories) suggested that CNCI for Sri Lanka was similar to that of the United States and Germany. The use of the ESI schema or either of the ANZ FoR schema produces an outcome in which Sri Lanka is apparently world-beating. Indonesia's CNCI is also elevated if these schema are used, but the CNCI of most countries is generally affected much less although that of the United States, United Kingdom, Australia and Germany are all slightly depressed under FoR Level 2 and the Polish PL-19 schema. Indonesia benefits under the Polish schema but Sri Lanka does not.

They key lesson here is that the way in which the data are selected and aggregated will have an influence on analysis and interpretation, yet none of these alternative schema have been implemented casually or without planning, analysis and prior development.

Collaboration

The global research landscape has changed considerably over the last forty years. In the 1980s it was dominated by a *trans*-Atlantic axis with links to Japan and to Anglophone countries with established university systems on the European model. In 2020, the balance of the research world has changed: Asia-Pacific plays a key role, through China (the second largest research economy in **Table 1**), South Korea, Singapore and a network that stretches to Australia (higher CNCI than the United Kingdom or United States in **Figure 1**); there is another, growing network across the Middle East and North Africa; and Latin America waxes and wanes as economic cycles create opportunity.

There has been an increasing level of international collaboration across this dynamic world network (Georghiou, 1998; Wagner and Leydesdorff, 2005, Wagner, 2008; Leydesdorff and Wagner, 2008). International collaboration has generally been seen in policy research discussion as a supportive research strategy enabling access to greater intellectual and

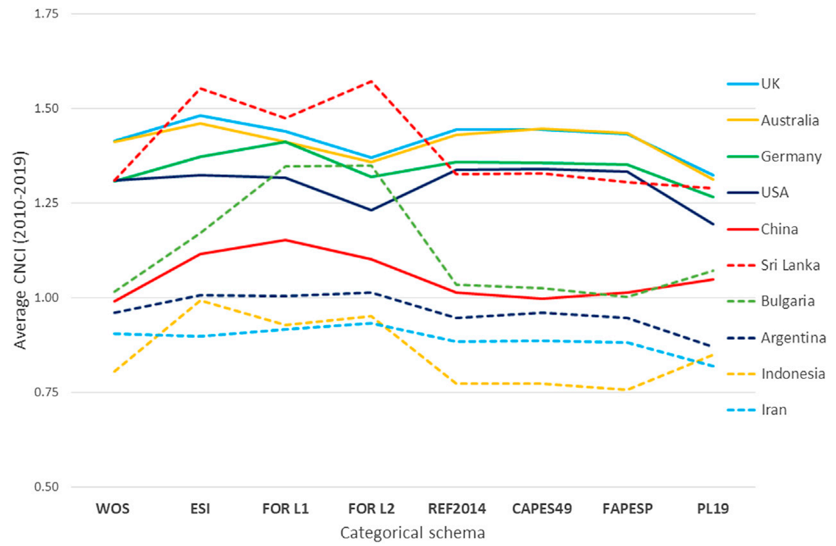

Figure 3 | The average Category Normalized Citation Impact (CNCI) for ten countries calculated with data normalized under eight different categorical schema. The numbers of publications used to calculate CNCI vary between schema as indicated in **Table 3**. The graph lines do not imply any connection between distinct schema but are inserted as a visual aid.

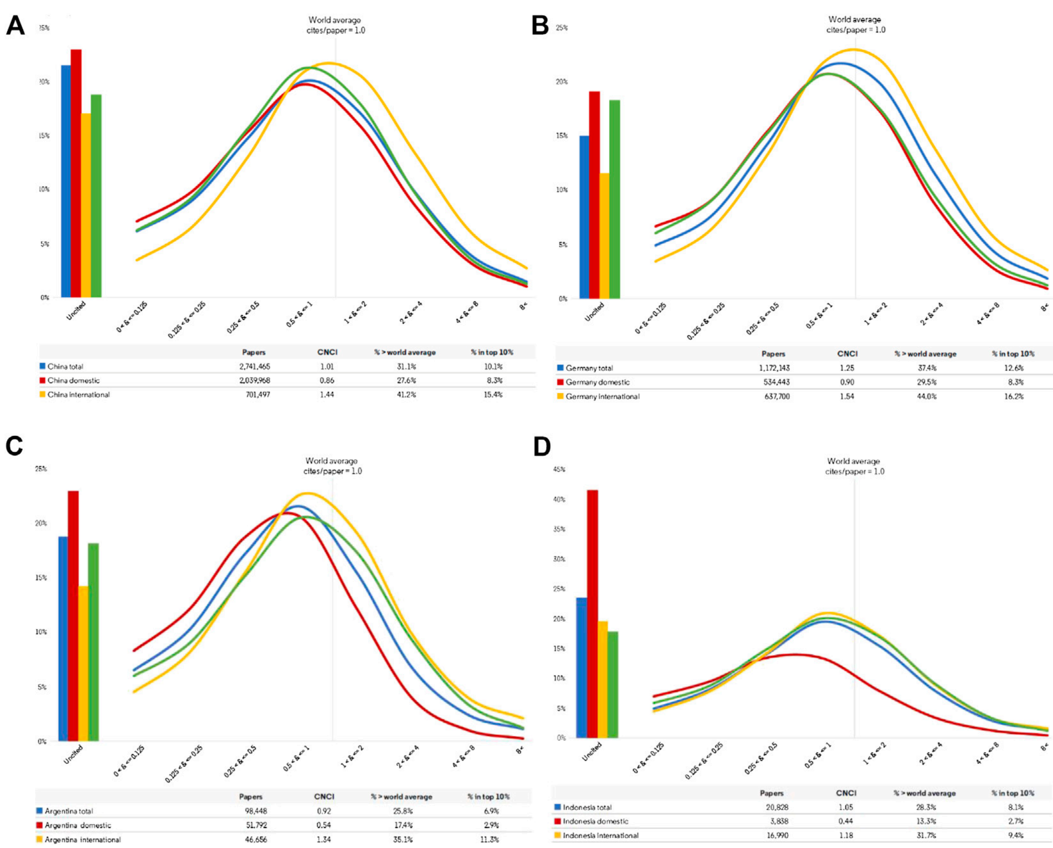

Figure 4 | (A)–(D) Impact Profiles for four G20 countries for articles and reviews indexed on the *Web of Science* during the ten-year period 2009–2018. Each profile includes three extracts for the country plus a reference benchmark taken from the complete G20 dataset. The three extracts for each country are the Impact Profile curves for: total national output; domestic output (with no international co-author); and internationally collaborative output. Each curve shows uncited papers (histograms to the left) and the distribution of output across eight categories of increasing impact relative to world average. The green line is a common reference set for all the graphs and marks the average for the complete G20 dataset. **(A)** China **(B)** Germany. **(C)** Argentina **(D)** Indonesia.

economic resources and accelerating work both on researcher-driven projects and on strategic programs such as those in particle physics and on the human genome. For this reason, it is often monitored and promoted as part of national research policy (for example, in EU policy and the EU's Horizon 2020 research program (https://ec.europa.eu/research/iscp/index.cfm?pg=policy). It is also associated with increasing citation impact (Persson et al., 2004) and internationally collaborative papers are more frequently cited on average (see **Figure 4** later).

Analyses by ISI (Adams, 2012; Adams, 2013) over the last ten years have identified changes consequent upon these increases in international collaboration that alter the structure of the national research base. For Australia and Germany, as examples of large research economies, international collaboration has in fact become the critical driver of rising productivity (**Figures 5A,B**). Almost all increase in annual publication counts can be accounted for by output shared with one or more collaborating countries whereas the domestic research output (with no international co-authors) has plateaued.

The pattern for countries that are still growing and developing their research economies may be quite different. Indonesia's overall output has risen steeply but its level of international collaboration has always been very high and has increased so that a very high proportion of its output over the last decade has been collaborative (**Figure 5C**). Iran also has steeply rising research output but it is almost entirely driven by the domestic research base and its international collaboration has been much lower (**Figure 5D**).

The United Kingdom and Germany share around 10% of their output with one another and each shares around two-thirds of its annual output with other countries. This pattern is similar across the European Research Area and mirrored by most other advanced economies. The internationally collaborative part of each country's output is also the more highly cited (Adams, 2013), which is unsurprizing since collaboration requires a shared agenda: a compromise that must be offset by clear likelihood of research benefit.

The innate, historical research strength of the larger, established economies countries means that while collaboration may boost their performance as measured by average CNCI it does not alter it disproportionately. However, the contribution made by different partners is not uniform. Adams and Gurney (2018) showed that the United Kingdom 'gained' in citation impact when collaborating with the United States, Germany and France and the average CNCI of such papers was as much as twice world average. This citation boost changed when, instead of all co-authored papers irrespective of third parties, only bilateral papers were considered. The United Kingdom still gained but for German and French collaborations it did so only marginally. This separation of bilateral and multilateral components may become increasingly important (see also **Table 4**).

Disproportionate change due to collaboration can compromise the research metrics of smaller economies such as Indonesia with a shorter history of investment and growth. We analyzed the parts of national output that are accounted for by domestic authorship (both single and multiple), bilateral international collaborations, trilateral and multilateral collaborations. We counted the numbers of articles and reviews produced over the decade from 2009 to 2018 and calculated the share of total citations attributable to each

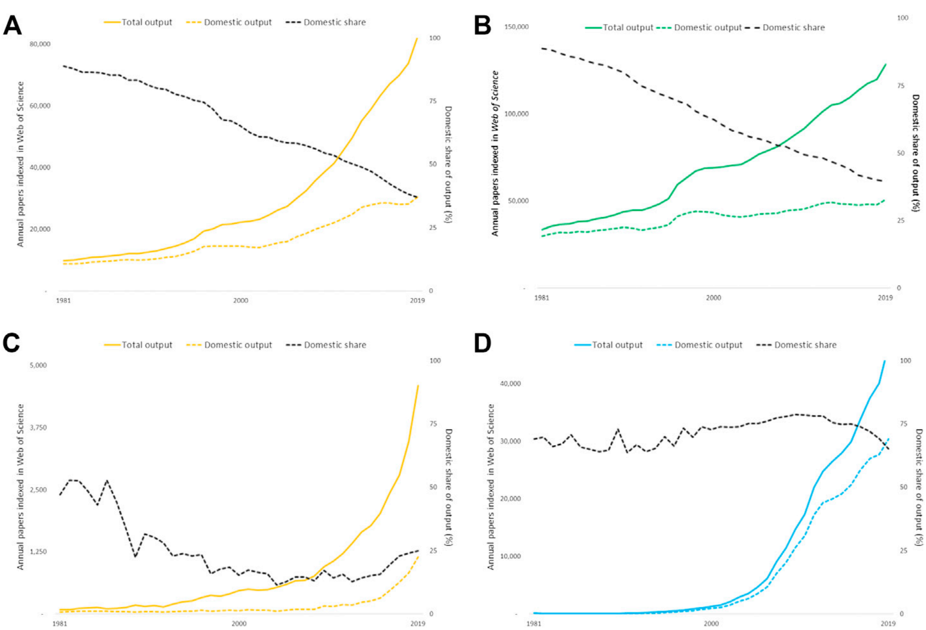

Figure 5 | (A)–(D) Output indexed on the *Web of Science* for Australia, Germany, Indonesia and Iran deconstructed by total and purely domestic articles and reviews. The domestic share of output has steadily declined for the large research economies while output, boosted by collaboration, has steadily risen. Output for the smaller economies has risen more steeply but the profile of international collaboration is less consistent. **(A)** Australia **(B)** Germany. **(C)** Indonesia **(D)** Iran.

TABLE 4 | Total national papers and those co-authored between a European country and former colony (2015–2019). Collaborative papers may have other, third-party countries as co-authors so both the total collaborative and the solely bilateral counts are shown.

Collaboration: All/bilateral		France	Netherlands	Spain	United Kingdom
	National total	398,747	221,375	321,566	666,166
Argentina	49,997	3,743/878	1,883/102	5,789/2,190	3,418/312
Indonesia	15,333	932/196	1,476/574	366/26	1,654/288
Kenya	10,842	720/64	890/120	400/11	2,783/521
Tunisia	23,013	6,973/4,670	186/9	1,547/795	596/87

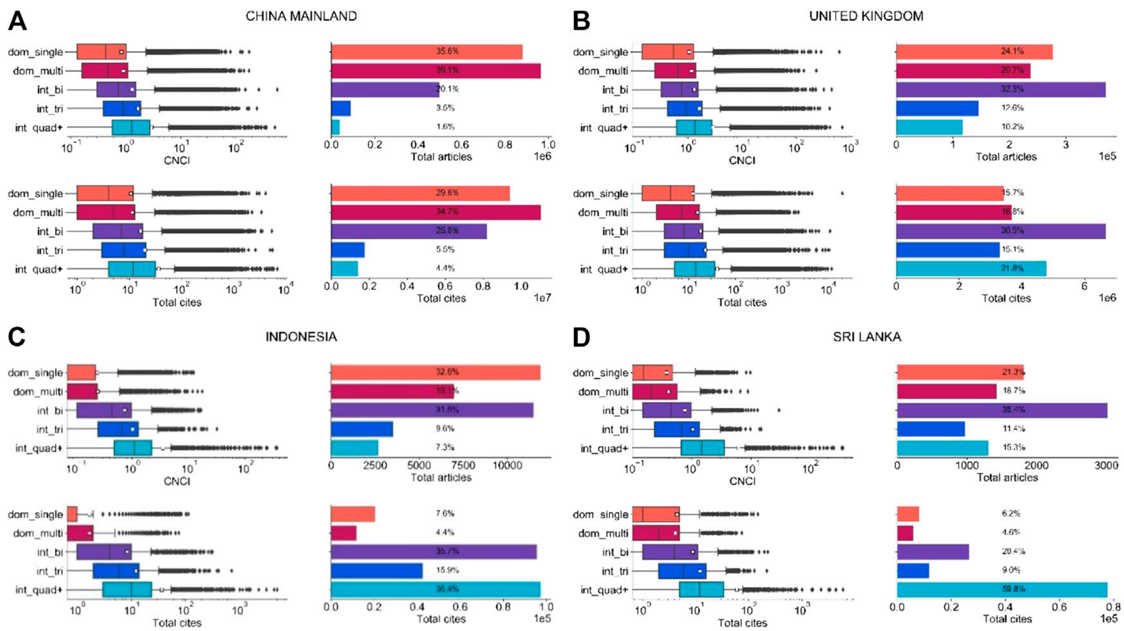

Figure 6 | (A)–(D) Deconstructed article and citation distribution output for China, the United Kingdom, Sri Lanka and Indonesia (2009–2018). Articles are divided into five types: domestic single (dom_single) and multi (dom_multi), and international bilateral (int_bi), trilateral (int_tri) and quadrilateral-plus (int_quad+). White squares on boxplots represent the mean. Data reproduced from Potter et al. (2020, **Figure 6**). **(A)** China **(B)** United Kingdom. **(C)** Indonesia **(D)** Sri Lanka.

country that were contributed by these different groups of papers (Potter et al., 2020). (**Figure 6**)

Domestic output for China is 75% and for the United Kingdom is 44% of total published output. This accounts for 65% of China's citations and 35% of the United Kingdom's (**Figures 6A,B**). Domestic output therefore collects *pro rata* a similar but smaller proportion of citations than it represents as a proportion of publications. However, domestic output for Indonesia (51.7% of total output; **Figure 6C**) and Sri Lanka (38%; **Figure 6D**) accounts for a much smaller share of national citations received (around 10%). By contrast, their highly multilateral papers (respectively 7.3% and 15.3%) account for, respectively, 36.4% and 59.8% of the citations they received (**Figure 3**) and, thus the overall CNCI figure is highly dependent on the performance of the multilateral collaborations to which they contribute. By contrast, highly multilateral papers were 1.6% of output and 4.4% of citations for China and 10.2% of output and 21.8% of citations for the United Kingdom. In other words, the large research economies not only gain relatively more citations from domestic output but while their multilateral collaborations certainly augment overall performance metrics they do so to a lesser extent: by a factor of 2 rather than four to five.

Referring back to **Table 1**, we conclude that the exceptional average CNCI for Sri Lanka appears to be dependent primarily on its collaborations rather than its innate research profile (**Figure 6D**) whereas that of China is clearly proportionate to the balance of domestic and collaborative activity (**Figure 6A**). The particular annual values are dependent on the numbers of such collaborative papers in that year and the time they have had to receive international recognition whereas the more stable CNCI metrics for the large economies are attributable to innate national research activity and recognition. The implication is that it is not sufficient to evaluate national bibliometric performance solely through summary indicators but it is essential to understand the balance and stability of domestic, collaborative and highly collaborative activity that feeds into such indicators and to be aware of which other countries may be involved in such collaboration (see also **Table 4**).

Fractional Attribution

It has historically been the practice to assign the full value (of both production credit and CNCI value) of a publication to each author, each institution and each country listed in the author metadata. This may cover participation but it does not necessarily reflect contribution. Given the collaborative nature of research, it has been argued that fair assignment of credit to the authors is not only important but essential (Allen et al., 2014) and this perspective is increasingly supported by the academic community because of its significance for funding (Sivertsen, 2016), promotion (Klein and Falk-Krzesinski, 2017), and national standing (Ahmadpoor and Jones, 2019). However, Larivière et al. (2020) both argue that the interpretation of contribution roles may vary as widely as criteria for authorship in different disciplines and that attribution of leadership and supporting roles may become a divisive and value-driven process.

One frequently proposed alternative is fractional counting (Waltman and van Eck, 2015) whereby each author is assigned part of the credit and CNCI value. From an aggregate perspective, fractional counts add up to the same number of articles as are in the data, which may provide better balance and consistency in bibliometric indicators but it is also claimed to improve precision: an assertion that is unprovable and misleading. Equal is not the same thing as equitable in the distribution of credit, and this is evident among international multilateral papers (**Figure 6**).

An even fraction may accurately reflect credit for some small groups (perhaps up to four individual entities?) but no algorithm will allocate credit proportionately among larger groups where major and minor contributors must be present. Sivertsen et al. (2019) showed that median authorship rates vary markedly between fields. They proposed a family of indicators for modified fractional counting (MFC) based on the root of the fractional authorship, which they argue eliminates extreme differences in contributions over time that otherwise occur between scientists that mainly publish alone or in small groups and those that publish with large groups of co-authors.

Another approach is to enhance CNCI normalization. There is a clear disparity in article volume, citations and CNCI between different collaboration types and countries (**Figure 6**). Potter et al. (2020) proposed a new metric, 'Collab-CNCI', that accounts for the level of collaboration without presuming credit. Their analysis demonstrates that Collab-CNCI reduces the impact of highly collaborative articles on a country's mean CNCI when using the full count method, providing a more balanced view than the standard mean CNCI. The relative decrease in mean CNCI was greater for the smaller research economies, where, generally, multilateral collaborations make up for a larger and sometimes disproportionate percentage of their publication output.

History and Geography

The collaborative links for many research economies are influenced not only by their capacity, but also by their geography and history, particularly where there are significant global links to former world powers.

The United States appears to be less collaborative internationally than other G7 economies (Adams, 2013) but this may be, at least in part, a consequence of its location (with borders on the Atlantic and Pacific Oceans) and the great size of its domestic economy. It is as far, and takes as long to fly, from Los Angeles to Boston as from London to Ankara but the latter route crosses many borders in the European Research Area. New Zealand's remote location may explain why it is less collaborative than the similarly sized Denmark: both are strong research economies but the latter is positioned in the European network.

Links to former colonial powers are also reflected in many concentrated collaborative partnerships. We can consider the relative number of collaborative papers between four large European research nations that previously occupied territories in other parts of the world. Comparison of total and purely bilateral international collaboration suggests that historical ties and language shared between Spain and Argentina make this a stand-out relationship for both countries. About 12% of Argentina's publication output is collaborative with Spain and more than one third of those papers are purely bilateral, with no third-party participation. This compares with its collaboration with France, the United Kingdom or the Netherlands where it has fewer shared and many fewer bilateral publications. France evidently has a far stronger relationship with Tunisia and collaborates on almost one-third of that country's publications, with a high proportion of purely bilateral co-authorships (**Table 4**).

The United Kingdom has strong ties to Kenya and is a co-author on about 25% of that country's papers, many more than any other EU nation. The five-year total tally is actually fewer than that between the United Kingdom and Argentina, but the bilateral tally is not. The significance of the bilateral component is again affirmed by the links between Indonesia and the Netherlands: the larger United Kingdom has slightly more collaborative papers with Indonesia than does the Netherlands but the latter has twice the number of bilateral co-publications.

The significance of these national links is that they are an indicator of two things: a prior cultural influence that is likely to be reflected in the research structure and portfolio of the growing economy; and an overlapping component in publication and citation data. It is infeasible, for example, that Tunisia's average CNCI is not associated to a marked degree with that of its collaborators in France.

The overall pattern of collaborative links for Africa (Adams et al., 2014) confirms the residual legacy of previous colonial links, often traceable to institutional associations through a shared European language that became the foundations for later collaborative networks. A West Africa group (Benin-Togo) pivots around Cameroon, a relatively research productive country, and the common factor within this group is almost certainly their common use of French as the cross-national business language. A large group of collaborative nations in East Africa includes Kenya and geographical neighbors but also includes West African Nigeria, Ghana and Gambia which share English as a common language.

Such bilateral connections and local networks, drawing on a history and investment beyond the global milieu, contribute positively to overall performance profiles. It is essential to be aware of such histories in interpreting and explaining activity and performance patterns for both the established and growing partners.

We also note again the effect of geography. This is immediately obvious in the East and West networks within Africa although they are also influenced by the major communication factor of shared language. In North Africa, we can see that the Mediterranean location of Tunisia has sustained its historical links to France and enable it also to have substantial collaboration with Spain.

Culture

The calculation of CNCI draws upon our understanding that citation counts not only grow over time but do so at rates that vary by discipline. They are influenced by disciplinary cultures: at a broad level, between humanities and the natural sciences; at an intermediate level, between organismal and molecular biology; and at a fine level, between basic and applied work on the same topic.

A further factor, that is less often identified or understood, is the influence of national cultural differences, influences that appear linked to perceptions of the relative significance of domestic and international research.

English has become the lingua franca of international research and the use of other languages impacts visibility and citation potential (van Leeuwen et al., 2000; van Leeuwen et al., 2001). For example, Russia and Brazil exhibit strong preferences for the Russian and Portuguese languages, respectively, even within journals indexed in WoS. The extent to which a nation's output appears in journals with a domestic rather than an international orientation appears also to have a subsequent effect on citation potential. Japan is an example of a nation that disproportionately publishes in the domestically oriented journals of the nation's scientific and medical societies. Even when these titles are English-language and published by international commercial firms, their content is less seen and less cited than papers appearing in internationally oriented journals (Pendlebury, 2020).

Another example of the influence of national, and likely cultural, factors on indicators of national research performance is seen in our analyses of data comparing the CNCI trajectory of China with the US and major European economies, which brings out a further example of misunderstanding what particular constructions of the citation data are reporting.

The data in **Figure 1** appear to present CNCI tracks for ten countries across a five-year period. In fact, the annual data points show the average citation count to date for the papers published in each of those years. The CNCI indicators for the papers published in 2015 are informed by five years of citation data, at both national and global benchmark level. The papers for 2019 have one year's accumulation of citations at best and much less for those papers published later in the year.

The format of **Figure 1** is typical of that in many national and agency reports, but is it a fair reflection of performance and, more specifically, of the trajectory of performance across a period? Concern about the number of analyses that appeared to suggest that China, despite its growing research investment, was failing to deliver research of quality, led ISI to an analysis that compared the picture presented by the traditional historical analysis with a different deconstruction, one that followed the performance of an annual cohort of papers as citations accumulated over time for specific countries and for the global benchmark: this presents a different perspective (Adams, 2018). The key here is that the indicator is not tracking a change in performance over time but the record of performance for different cohorts of papers based on citations recorded to date.

In conventional analysis, analysts illustrate a performance trend using all available data, which means counting and normalizing all citations to date for the publications of each year in the series. The series then shown is not a performance track for any particular set of papers but an implied track for the entity as a whole, where normalization compares the entity to the global average. A conventional time-series analysis based on all the available data at a single census point (drawn from the National Science Indicators published annually from 1992 by ISI and then by Thomson Reuters) would, for example, suggest that Germany is on a clear upward trajectory but that while the average CNCI of China's output is unquestionably improving, it tends to fall in relative performance in the most recent year of each series (**Figure 7**).

However if, instead of looking at the CNCI for publications in a series of years, we track papers from a particular year over time as citations accumulate both to our target cohort to the rest of the world's papers published in that year, then we see that the CNCI of German papers falls in later years after a relatively high level achievement in the years immediately after publication (i.e., the 2006 papers have their highest CNCI in the 2007 series and then drop lower in each later version). Each time series in successive versions of our NSI versions essentially mimics that of the previous and there is little net improvement. By contrast, the trajectories for China progressively improve in CNCI relative to world average (i.e. the penultimate year of every series is at a successively higher CNCI value than any previous publication set).

Annual United Kingdom CNCI data follow the same pattern as Germany and the United States falls off even more markedly. Which is the 'correct' analysis? Neither: both are necessary for a fuller understanding of performance dynamics.

Thus, it appears to be China and not Germany which is 'on the up'. Why should the citation impact trajectory of China's output differ from that of the West? We cannot be certain about this but there are several possibilities. First, there may be a tendency in Western research economies to focus on 'recency' where the latest research garners particular attention. The publications of the most recent years are those frequently cited and the citation count plateaus rapidly after that initial burst of attention. By contrast, the rapidly expanding output of Chinese researchers may be

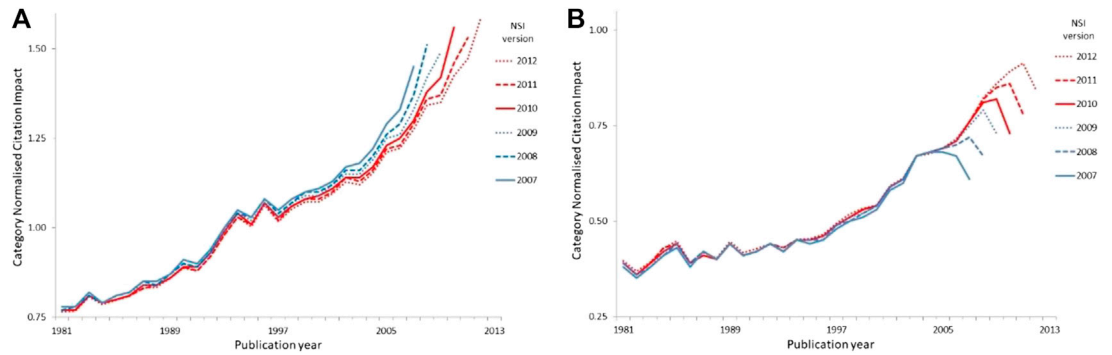

Figure 7 | (A), (B) Average annual CNCI for Germany and for China as captured in successive time series from National Science Indicators (NSI, using versions published annually from 2007 to 2012) based on data in *Web of Science*. Each successive NSI time series uses all available data at the census date to illustrate an implicit national trend based on citation counts to date for the publications of each year, with the publication year indicated on the *X*-axis. Comparison vertically, i.e. between the NSI values of a specific publication year, shows the real CNCI trend of any one year's papers in successive NSI versions. **(A)** Germany **(B)** China.

referencing the smaller body of slightly older literature which then boosts the relative citation status for those cohorts. Thus, after five years or so the average CNCI for Chinese literature has moved up on the world average while the German, United Kingdom and US literature has dropped back. A second possibility is that Western literature retains a primacy while China is still establishing its global profile. Thus, both Chinese and Western researchers focus on the latest discoveries in Europe and America first and then only subsequently does the Chinese research base recognize its own achievements.

Tang et al. (2015) have drawn attention to a "clubbing" effect in China's recent surge in research citations. For highly cited nanotechnology papers, they found that a larger proportion of Chinese citations are from domestic institutional and national networks than is true for similar U.S. papers. This may be a cultural factor, but it may equally be an indication of the degree to which Chinese nanotechnology research, which has grown to twice the size of the US, is now more citable.

Clearly, context must be assessed as well as data. Whatever the explanation, the key effect on the interpretation of research metrics is that performance trends need careful interpretation in a full understanding of the basis on which a time series has been analyzed.

Global Benchmarks

Another, apparently artifactual and potentially confusing outcome of the pervasive growth of collaboration is that it is possible for all countries to have a CNCI value that is above the world average and yet to have more than half their output below world average. This contextual information is rarely apparent to subject-expert evaluators and may consequently be disturbing when encountered.

The explanation is that the global total must include all the national pools of domestic papers (relatively less often cited) plus a single, deduplicated set of the shared pool of internationally collaborative papers (on average more highly cited). By contrast, each country has only its own pool of domestic papers plus its portion of the collaborative pool.

This may still seem infeasible but the schematic analysis in **Figure 8** for a hypothetical world of three small countries shows that the global benchmark can indeed be below all three of the contributing nations' individual citation averages.

It is equally the case, for a country with an average CNCI above 'world average', that more than half of the country's papers will have individual CNCI below world average. The initial reaction of research managers will be that this is not possible but it is in practice not only possible but a likely consequence of the skewed nature of citation distributions that result in an average value that is well above the median. Many papers in most samples are uncited, possibly because they are recently published; most have a modest number of citations; a few will have attracted many citations. This skew is familiar to scientometricians but not to research-domain specialists and it leads us to the need for graphic illustrations of the distribution of impact that underpins the averages.

A problem arose in reality when ISI was faced with two apparently similar biomedical research units under quinquennial review which appeared to have very different performance as indicated by their average CNCI (the report on this is commercially sensitive). The solution to improved understanding, and the route to a graphical analysis that would inform and support management decision making, was to visualize the distribution of performance in 'bins' ranked by relative citation performance around the world average. By separating out the frequent uncited papers and then ranging the remainder in eight tranches with successive doubling of their relative impact, it is very easy to see the shape of the distribution, the balance of exceptional and weak research and to compare multiple curves or 'Impact Profiles' (Adams et al., 2007). In the particular instance that drove this development, it became evident that a very small number of exceptionally highly cited papers for one unit strongly skewed, even 'distorted', its average but the overall Impact Profiles were otherwise identical. The analysis thus validated the original views of the expert review group.

The Impact Profiles all confirm the influence on citation impact of the internationally collaborative component of each

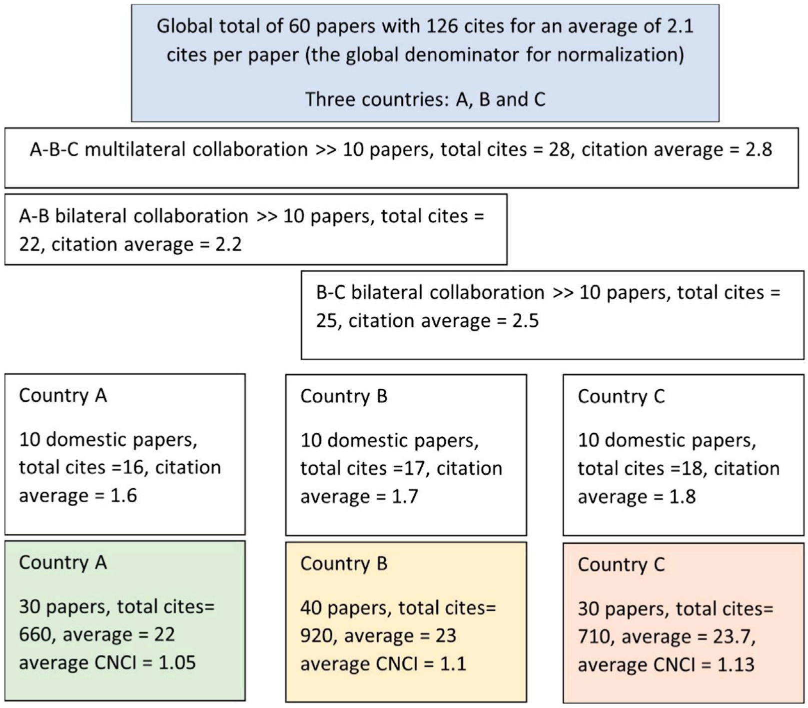

Figure 8 | For a hypothetical universe of three small countries, with some shared and some purely domestic papers, it is evident that the global average citation count may be less than that of any one contributor. This is because domestic papers are on average less frequently cited than internationally collaborative work and the global total includes all three domestic pools whereas each country only hosts its own.

country's activity. They show that the CNCI distribution is almost always spread across a range of impact categories from well below world average, where CNCI is 1/8th or less of world average, up through successively higher tranches. Similarly, while Germany has a high average CNCI (**Table 1**), it still has a substantial output of poorly cited papers, which may be a language effect (**Figure 4B**). No country is either completely excellent or uniformly poor in its research. Impact Profiles also enable us to introduce a reference curve, not just a single metric such as 'world average' but a complete profile for either the world or, as in **Figure 4**, a relevant reference group, which is the average for the combined G20 dataset. This also enables rapid comparison between the different countries.

An important aspect of the Impact Profile is, therefore, that it not only properly presents the distribution underpinning the CNCI indicator but it also reveals the extent to which a country (or institution or group) that has only a modest average impact may nonetheless have excellent papers in its portfolio. Furthermore, it establishes a much better contextual comparison because it does not use a single point metric as a benchmark but it can deploy a reference curve across an entire distribution. This has immediate practical applications in any research evaluation since the appearance of the higher impact papers in a profile will then prompt management questions about their authorship, the source of their citations and their links to–perhaps even dependency on - other, less prominent work. Research development and investment is facilitated by moving away from a summary to unpack the content and see a route to action.

Context and Distributions

A shift from CNCI toward a more contextual basis for analyzing citation counts has been advocated by scientometricians (e.g., Waltman and van Eck, 2016) who have pointed to the value of percentiles as a tool for moderating both skew and kurtosis in citation distributions. The latter means that in some low-citing fields it would be exceptional to have a paper that was much above four times world average whereas in fields of citation abundance the greater spread of counts facilitates values more than eight times world average.

Bornmann et al. (2012) point to the use of a percentiles as an improved basis for an indicator of excellence in world rankings and Bornmann (2013) highlighted their analytical use in research evaluation, enabling both an assessment of the distribution of percentiles across a set and a focus on the publications with the highest citation impact. Waltman et al. (2012) discuss possible statistical problems in ranking caused by the discrete nature of citation distributions, especially with small samples, and applied a fractional solution. Bornmann and Williams (2020) discuss this and elaborated on earlier work to describe guidelines and procedures for the normalization of percentile ranks based on cumulative frequencies in percentages. They also show how graphical visualization can present this information in a more meaningful and accessible manner.

Although we have encountered an interpretive problem, in that percentiles suffer from a lack of intuitive understanding among casual users, and they may also be unsatisfactory with small samples, we nonetheless agree that percentiles generally provide a better explanatory context than CNCI for understanding the impact of a paper in its field. We note, for example, the methodology used in the Leiden Ranking of world universities (https://www.leidenranking.com/ranking/2020/list). This ranking draws on percentiles rather than normalized citation counts and applies a threshold at the top 10% of papers by field, ranking institutions according to the overall proportion of their papers that pass such a field threshold (Waltman et al., 2012).

Context and Maps

We noted at the outset that contemporary bibliometrics can go further and address other contextual criteria set out by the ABRC (1987) including timeliness and pervasiveness. While percentiles clarify relative excellence, they do not increase the evaluators' understanding of significance in other contexts. To do this it is necessary to determine whether the research under evaluation is part of current and substantive developments in its field, or in associated fields where it has application. Is it a part of a research cluster that is currently well-cited (timely) and is that cluster significant in scale and reach (pervasive)?

In developing the Science Citation Index, Garfield (1955) recognized that citation data provide material to build a picture of the structure of scientific research and sketch its terrain. In the previous section we arrived at Impact Profiles, which enable us to see the distribution of excellence in any dataset and then set that against a reference curve that lifts our appreciation of context beyond a point metric such as world average. This is a statistical relationship. In addition, once an index linking papers through their citations exists, there is a basis for determining their intellectual relationships. Derek de Solla Price (1965) noted, "The pattern of bibliographic references indicates the nature of the scientific research front." This pattern provides a map in which a research publication can be located and from this the analyst can apply a time axis that shows the direction of intellectual travel. It is possible to determine where a topic is and what direction the research around that topic is taking.

Small (1973) laid the foundations for defining specialties in research fronts using co-citation analysis. Small and Griffith (1974) and Griffith et al. (1974) showed that individual research fronts could be measured for their similarity with one another and thus form the nucleus of a specialty. Their mapping used multidimensional scaling and similarity was plotted as proximity in two dimensions. There are now many academic centers across the globe focusing on science mapping, using a wide variety of techniques and tools (Börner, 2010; Boyack and Klavans, 2010; Petrovich, 2020). These later developments are summarized in Indiana University Professor Katy Börner's (2010) *Atlas of Science*. Of particular significance are CiteSpace developed by Chen (2006) and VOSviewer developed by Van Eck and Waltman (2010) at CWTS, Leiden University.

The approach to mapping scientific and scholarly research as traditionally employed at ISI and devised by Small is as follows: A research front appears when a set of recent publications all co-cite several earlier papers that stand out because they are themselves in the top 1% (the highest percentile class) for their year and field. The recent papers are linked by the highly-cited targets they cite in common and thus form an emerging front of research activity, the identification of which may be determined by a review of their common keywords (**Figure 9**).

For a research evaluator the first question is whether the work that they are reviewing appears in one or more of these research fronts. They can then use research fronts to address their knowledge of the additional issues of timeliness, which may be determined by the recency of the citing papers, and pervasiveness, which may be inferred by citation abundance and spread across fields.

More generally, for an institution, how much of its work is in or (extending the mapping analysis) close to a research front? Important management opportunities, which go far beyond the information derived from research performance metrics, appear when research fronts are precisely located in the knowledge network. A research manager can determine the distribution of institutional output across the knowledge landscape, filtering for recent or longer time windows, and then assess the relationship of their research clusters to a front. They can also make a comparative evaluation with competitor institutions. Similarly, research funders, by identifying the distribution of publications arising from funded projects, can see whether investments are producing work located in or near research fronts and policy makers can use this approach to map research emerging at a national level (Chinese Academy of Sciences, 2019; Igami and Saka, 2016).

Context and Purpose

We started by noting that research evaluation is usually interested in excellence (Moore et al., 2017; Ferretti et al., 2018) and that "excellence depends on context" (Nature, 2018). The reach of and attention given to an innovation in an emerging cross-disciplinary research area will be very different to research with the direct but narrow impact critical to solving a technological constraint for an industrial process. What is true is that in both instances the research will only be 'good, valid, timely and useful' if it is high quality, yet that quality will not be measured by stakeholders in the same way.

Intention, purpose and objectives should be an embedded component of the initial design of every research assessment process. Why are we doing this, what do we seek to discover, what would tell us whether this research is good and what tells us whether it has achieved its aims? If an assessment starts without these criteria in mind and without adapting and matching the data, methods, analysis and indicators to those criteria then it is less likely to provide a satisfactory and informative interpretation of outcomes for the user.

We refer again to the perspectives listed by Langfeldt al (2020) and their relevance to the ABRC (1987) internal and external criteria. The values of novelty and utility are not the same thing. Both require 'good' research but the index of goodness for one

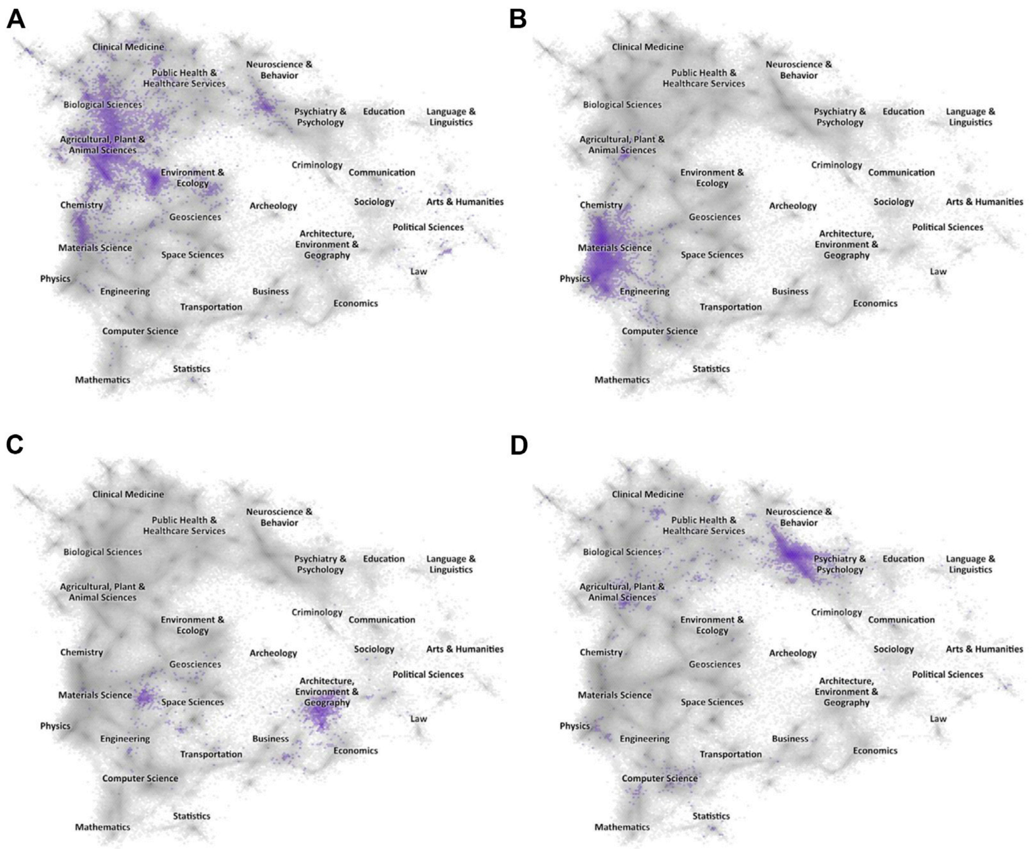

Figure 9 | (A)–(D). Science maps showing the spread of articles in four topical Research Fronts. The greyed landscape is the proximity map for all articles indexed in the Web of Science (2010–19) within which major domains have been labeled, and the highlighted area in each map are those papers linked to a specific topical Research Front. **(A)** CRISPR **(B)** 2D Materials. **(C)** Global Energy System Transition **(D)** Dynamic Functional Connectivity.

may not be consonant with the other. Similarly, the value dimensions of researchers, research funders and national research policy will be conditioned by the objectives peculiar to each. The suitability of any bibliometric approach is proven by defining those objectives and setting the analysis in a structure that matches method to purpose.

Discussion: Implications for the Original Example

There is a need for changes in the approach to using bibliometric data: the subject-expert user needs to be clear whether the data they have are relevant to the evaluation questions they pose; they need to establish an *a priori* understanding of how they will use the data and of the choices of methods to apply; and it should be standard practice that data are developed and presented not as summary point metrics but in a form that allows accessible, interpretive exploration through drilling down or 'peeling the onion' of any rich analysis.

It should be clear from this review of scientometric data underpinning bibliometric indicators that, when looking back at **Table 1**, an evaluator would be incautious if they were to rely solely on summary information to make judgments about the relative or absolute research strengths, even of whole countries. This should be even more true if they were reviewing a table of institutions from the same countries or a set of their research groups seeking funding, and yet this happens frequently.

Highly granular categorical systems group research papers into small, self-referential pockets that boost the apparent relative citation performance of work which appears poorly cited in familiar topical aggregations (**Table 2**). More generally, the effect of a choice of discipline/topic categories for aggregating publications and normalizing citations is two-fold. First, countries with a less developed domestic research base, and less well cited domestic research output, will tend to have smaller publication tallies when more exclusive categorical systems (such as ESI and the ANZSRC FORs) are used (**Table 3**). Second, because such categories focus on journals selectively, it is the least well cited part of a country's activity that is omitted, so their average CNCI is raised (**Figure 3**). So, although publication counts for Sri Lanka, Bulgaria and Indonesia are significantly reduced in an ESI analysis compared with a WoS analysis they nonetheless then have higher average CNCI.

International collaboration is a pervasive factor for all countries and may cover much more than half their annual publication output, but the situation for smaller research economies is diverse (**Figure 5**). These collaborative papers are more highly cited on average, for all countries, and thus raise their average CNCI. For smaller countries, the balance of output and citations becomes disproportionate: for Indonesia 52% of papers are domestic but 88% citations come from international collaborative papers; for Sri Lanka the figures are 34% domestic papers and 90% international citations. Iran, by contrast, relies largely on its domestic research output. In consequence, **Table 1** should be re-interpreted in the light of the balance of domestic and collaborative output and citations in each portfolio, and **Figure 4** further emphasizes the potential benefit due to collaboration as compared to domestic activity.

Historical links to well established European research economies can have a significant research benefit because of sustained collaborative partnerships. This is an excellent outcome at a cultural and economic level but it could be a covert factor influencing outcomes at a bibliometric level. Argentina's relationship with Spain and Indonesia's relationship with the Netherlands are examples (**Table 4**).

Cultural factors are rarely identified as a research analytical factor at national level, although they are widely acknowledged at a gross (arts/science) and fine (molecular vs. organismal biology) disciplinary level. The beneficial effect of 'recency' on citation rates for Germany (and other G7 research economies) is apparent in comparison with China, which appears to cite later but then to have rising relative citation performance for any year (**Figure 7**). This highlights the need to be wary of any short windows in an analysis, or of focusing unduly on the most recent data, without understanding the research culture and behavior of the target under analysis.

A further complication with international collaboration and the relatively higher citation counts for international publications (**Figure 4**) is the consequent effect on net national CNCI. Every national portfolio is enhanced compared to the global pool because it contains only the national slice of lower cited domestic activity. The best way to interpret the real distribution of CNCI is through a graphical analysis that reveals the full profile, the balance of work above and below world average, and the components due to domestic and collaborative output. Ideally, this would include a relevant benchmark.

We wholly endorse the views of Moed (2020) regarding the need for an evaluation framework in which the context and the purpose of the exercise are over-riding considerations. Citations are themselves value-laden constructs with social as well as research weight. Any aggregation of citation counts, subsequent management of the data through normalization and fractionation, and choice of analytical methodology then applied, must introduce further subjective modification that moves from original information toward a stylized indicator. The reader is referred to Ferretti et al. (2018) for a discussion of the challenge in establishing consensus on indicators of excellence.

In summary, the points that we have reviewed and of which those users planning a research evaluation should be aware are:

Normalization, granularity: a choice of broad or narrow focus is made when citation counts are normalized against a global benchmark, for comparative purposes or to aggregate data across years and disciplines (**Table 2**).

- USERS need to be aware of granularity and choose an appropriate level of aggregation.

Normalization, categories: there are many systems for assigning journals and/or individual publications to discipline categories and none is uniquely correct (**Table 3**, **Figure 3**).

- USERS should take the assessee's output portfolio into account in choosing a data source

Collaboration, domestic: the balance of domestic and internationally co-authored publications in a portfolio is likely to influence the evaluation outcome.

- USERS should be aware that papers with only domestic authors may be cited less often

Collaboration, impact: since internationally collaborative papers tend to higher citation impact the evaluator must reflect on the extent to which the data are driven by the target of evaluation or by work with its partners (**Figure 6**).

- USERS should consider the absolute and relative volume of international research collaboration

Collaboration, fractional attribution: it is argued that partitioning of credit for output and impact should be used to account for collaborative influence, but arithmetic solutions do not provably deliver greater precision or accuracy and are unlikely to assign the most appropriate fraction.

- USERS should be conscious of the balance of author counts in the evaluated output, and be aware of the effect of fractional attribution

History, legacy partners: the continuing influence of previous colonial relationships is evident (**Table 4**).

- USERS should recognize the legacy of history and consider how this might influence outcomes

Geography, distance and networks: not all countries are equal in their access to research partners by both distance and location.

- USERS should consider whether location factors may favor or constrain the assessed activity

Culture and language: there is a preference in some countries, sometimes stimulated by national Academies, to publish in nationally oriented journals and this, while entirely appropriate, naturally reduces exposure to external researchers who focus on 'international' journals.

- USERS should review the language balance in assessed output and any preference for journals with national rather than international orientation

Culture, national and disciplinary: differences in publishing and citing practice are known to exist between disciplines but it is less commonly acknowledged that distinctions in research culture also occur between countries (**Figure 7**).

- USERS should reflect on national and cultural components in data and indicators

Benchmark: the apparent anomaly that all nations can be above world average throws further light on the interpretation of trajectories (**Figure 8**).

- USERS must be sensitive to characteristics of the data and the analytical methods

Profiles: visualizing the full CNCI distribution in an Impact Profile not only shows the true spread of strong and weak performance around the average but also exposes the difference between that average and the median (**Figure 4**).

- USERS should seek data analyses that display the full distribution, not just point metrics

Context: most research indicators focus on a dataset for a target entity (country, institution, group) and the identification of research excellence. Research activity around the margins of that target and information in regard to other assessment criteria is less clear but it may be critical to interpretation and to the success of any intervention (**Figure 9**).

- USERS should consider that the research they assess is part of an ecosystem

The basic challenge for scientometrics is not about additional, new indicators but about presenting the outcomes of sound academic research in metrics and analytics in a form that domain specialist users can make use of for evaluation within their field. The future for the scientometrician should be less about the academic ideal in metrics, and its chimeric perfection, and more about user support including better management interpretation and faster, more confident decision making.

When the evaluator is clear about their objectives, the questions to be addressed, the relevance of bibliometrics to those questions and the nature of the available data, and the place of the bibliometric analysis within an overall evaluative framework, then they should proceed to work through the issues we list here and determine whether they have fully understood the implications of these and the outcome in the context of their purpose and materials. To facilitate such comprehension, this interpretation is preferably implemented locally, by the users (policy, funder, etc) and domain experts, rather than by an external analyst. The information presented must draw on a substantial body of data and may be best deployed not as tables but visualisations. It may also be that an intermediary - normally the secretariat supporting the decision-making group - is still required to mediate the interpretation. But this should now locate the target activity more closely for the evaluating group and in a meaningful context drawing on references to a wider information base that includes points familiar to multiple group members.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: Original data were sourced where indicated from the Web of Science, which is accessible to academic researchers in the United Kingdom under licence from the Joint Information Services Committee and in other countries through separate licensing agreements. Requests to access these datasets should be directed to https://clarivate.com/webofsciencegroup/solutions/web-of-science/contact-us/.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

All funding for this work was provided by the authors' employer.

ACKNOWLEDGMENTS

We thank colleagues not only at Clarivate but also at many university and research organisations around the world for their advice, contributions and many discussions during the analyses that preceded the writing of this article. We are also grateful for the feedback from two reviewers.

REFERENCES

Abramo, G., and D'Angelo, C. A. (2015). The VQR, Italy's second national research assessment: methodological failures and ranking distortions. *J. Assoc. Infor. Sci. Tech.* 66 (11), 2202–2214. doi:10.1002/asi.23323

ABRC (1987). "A strategy for the science base," in *A discussion document prepared for the secretary of state for education and science*. London, United Kingdom: HMSO, 1–50.

ABRC (1983). *The support given by research Councils for in-house and university research – report of a working party of the advisory board for the research Councils (the morris report)*. London, United Kingdom: HMSO.

Adams, J. (2012). Collaborations: the rise of research networks. *Nature* 490, 335–336. doi:10.1038/490335a

Adams, J. (1998). Benchmarking international research. *Nature* 396, 615–618. doi:10.1038/25219

Adams, J., and Gurney, K. A. (2018). Bilateral and multilateral coauthorship and citation impact: patterns in UK and US international collaboration. *Front. Res. Metr. Anal.* 3, 12. doi:10.3389/frma.2018.00012

Adams, J., Gurney, K. A., Hook, D. W., and Leydesdorff, L. (2014). International collaboration clusters in Africa. *Scientometrics* 98, 547–556. doi:10.1007/s11192-013-1060-2

Adams, J., Gurney, K. A., and Jackson, L. (2008). Calibrating the zoom – a test of Zitt's hypothesis. *Scientometrics* 75, 81–95. doi:10.1007/s11192-007-1832-7

Adams, J., Gurney, K. A., and Marshall, S. (2007). Profiling citation impact: a new methodology. *Scientometrics* 72, 325–344. doi:10.1007/s11192-007-1696-x

Adams, J. (2018). Information and misinformation in bibliometric time-trend analysis. *J. Informet.* 12, 1063–1071. doi:10.1016/j.joi.2018.08.009

Adams, J. (2013). The fourth age of research. *Nature* 497, 557–560. doi:10.1038/497557a

Ahmadpoor, M., and Jones, B. F. (2019). Decoding team and individual impact in science and invention. *Proc. Natl. Acad. Sci. U. S. A.* 116 (28), 13885–13890. doi:10.1073/pnas.1812341116

Aksnes, D. W., Langfeldt, L., and Wouters, P. (2019). Citations, citation indicators, and research quality: an overview of basic concepts and theories. *Sage Open* 9, 215824401982957. doi:10.1177/2158244019829575

Allen, E., Scott, J., Brand, A., Hlava, M., and Altman, M. (2014). Publishing: credit where credit is due. *Nature* 508, 312–313. doi:10.1038/508312a

ARC (2019). 2018-19. ERA national report. Available at: https://dataportal.arc.gov.au/ERA/NationalReport/2018/ (Accessed September 17, 2020).

Börner, K. (2010). *Atlas of science – visualizing what we know*. Cambridge MA: MIT Press.

Bornmann, L., de Moya Anegón, F., and Leydesdorff, L. (2012). The new excellence indicator in the world report of the SCImago institutions rankings 2011. *J. Informet.* 6 (2), 333–335. doi:10.1016/j.joi.2011.11.006

Bornmann, L. (2014). How are excellent (highly cited) papers defined in bibliometrics? A quantitative analysis of the literature. *Res. Eval.* 23 (2), 166–173. doi:10.1093/reseval/rvu002

Bornmann, L. (2013). How to analyze percentile citation impact data meaningfully in bibliometrics: the statistical analysis of distributions, percentile rank classes, and top-cited papers. *J. Am. Soc. Inf. Sci. Technol.* 64 (3), 587–595. doi:10.1002/asi.22792

Bornmann, L., and Marewski, J. N. (2019). Heuristics as conceptual lens for understanding and studying the usage of bibliometrics in research evaluation. *Scientometrics* 120 (2), 419–459. doi:10.1007/s11192-019-03018-x

Bornmann, L., and Williams, R. (2020). An evaluation of percentile measures of citation impact, and a proposal for making them better. *Scientometrics* 124, 1457–1478. doi:10.1007/s11192-020-03512-7

Boyack, K. W., and Klavans, R. (2010). Co-citation analysis, bibliographic coupling and direct citation: which citation approach represents the research front most accurately?. *J. Am. Soc. Inf. Sci. Technol.* 61, 2389–2404. doi:10.1002/asi.21419

Buckle, R.A., and Creedy, J. (2019). An evaluation of metrics used by the Performance-based Research Fund process in New Zealand, New Zealand. Economic Papers, 53 (3), 270–287. doi:10.1080/00779954.2018.1480054

Chen, C. (2006). CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature. *J. Am. Soc. Inf. Sci. Technol.* 57, 359–377. doi:10.1002/asi.20317

Chinese Academy of Sciences (2019). Research fronts: an annual report from the institutes of science and development. Available at: http://english.casisd.cn/research/rp/201911/P020191127531347988347.pdf (Accessed October 8, 2020).

CIVR (2006). Guidelines for research evaluation. Available at: http://vtr2006.cineca.it/documenti/linee_guida_EN.pdf.

DORA (2012). Declaration on research assessment. Available at: https://sfdora.org/.

European Science Foundation (2012). Evaluation in research and research funding organisations: European practices. A report by the ESF member organisation forum on evaluation of publicly funded research. Available at: https://www.esf.org/fileadmin/user_upload/esf/MO_Evaluation-RFO-Report_2012.pdf.

Evidence (2009). "Pilot study of bibliometric indicators of research quality: development of a bibliographic database," in *Report to the UK higher education funding bodies by evidence Ltd.* (Bristol, United Kingdom: Higher Education Funding Council for England), 1–83.

Evidence (2007). The use of bibliometrics to measure research quality in UK higher education institutions. Available at: https://dera.ioe.ac.uk//26316/ (Accessed September 17, 2020).

Ferretti, F., Guimaraes Pereira, A., Vertesy, D., and Hardeman, S. (2018). Research excellence indicators: time to reimagine the 'making of'? *Sci. Publ. Pol.* 45 (5), 731–741. doi:10.1093/scipol/scy007

Garfield, E. (1955). Citation indexes for science. A new dimension in documentation through association of ideas. *Science* 122, 108–111.

Georghiou, L. (1998). Global cooperation in research. *Res. Pol.* 27, 611–626. doi:10.1016/S0048-7333(98)00054-7

Georghiou, L. (1995). Research evaluation in European national science and technology systems. *Res. Eval.* 5, 3–10. doi:10.1093/rev/5.1.3

Gibbons, M., and Georghiou, L. (1987). *Evaluation of research, A selection of current practices*. Paris, France): OECD.

Glanzel, W., and Moed, H. F. (2002). Journal impact measures in bibliometric research. *Scientometrics.* 53, 171–193. doi:10.1023/A:1014848323806

Griffith, B. C., Small, H. G., Stonehill, J. A., and Dey, S. (1974). Structure of scientific literatures. II: toward a macrostructure and microstructure for science. *Sci. Stud.* 4, 339–365. doi:10.1177/030631277400400402

Hessels, L. K., and van Lente, H. (2008). Re-thinking new knowledge production: a literature review and a research agenda. *Res. Pol.* 37, 740–760. doi:10.1016/j.respol.2008.01.008

Hicks, D. (2010). Overview of models of performance-based research funding systems. Available at: http://works.bepress.com/diana_hicks/24/.

Hicks, D., and Melkers, J. (2012). "Bibliometrics as a tool for research evaluation," in *Handbook on the theory and practice of program evaluation*. Editors A. N. Link, N. S. Edward, and E. Vonortas (Cambridge, United Kingdom: Oxford), 323–349.

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., and Rafols, I. (2015). Bibliometrics: the leiden manifesto for research metrics. *Nature* 520 (7548), 429–431. doi:10.1038/520429a

Hirst, G. (1978). Discipline Impact factor: a method for determining core journal lists. *J. Am. Soc. Inf. Sci.* 29, 171–172. doi:10.1002/asi.4630290403

Igami, M., and Saka, A. (2016). Decreasing diversity in Japanese science, evidence from in-depth analyses of science maps. *Scientometrics* 106, 383–403. doi:10.1007/s11192-015-1648-9

Jappe, A., David Pithan, D., and Heinze, T. (2018). Does bibliometric research confer legitimacy to research assessment practice? A sociological study of reputational control, 1972-2016. *PloS One* 13, e0199031. doi:10.1371/journal.pone.0199031

Jappe, A. (2020). Professional standards in bibliometric research evaluation? A meta-evaluation of European assessment practice 2005–2019. *PloS One* 15, e0231735. doi:10.1371/journal.pone.0231735

Jiménez-Contreras, E., de Moya Anegón, F., and Delgado López-Cózar, E. (2003). The evaluation of research activity in Spain. The impact of the national commission for the evaluation of research activity (CNEAI). *Res. Pol.* 32, 123–142. doi:10.1016/S0048-7333(02)00008-2

Karolinska Institute (2010). External Research Assessment (ERA). Stockholm, Sweden: Karolinska Institute.

Ketcham, C. M., and Crawford, J. M. (2007). The impact of review articles. *Lab. Invest.* 87, 1174–1185. doi:10.1038/labinvest.3700688

Klein, J. T., and Falk-Krzesinski, H. (2017). Interdisciplinary and collaborative work: framing promotion and tenure practices and policies. *Res. Pol.* 46, 1055–1061. doi:10.1016/j.respol.2017.03.001

Lahtinen, E., Koskinen-Ollonqvist, P., Rouvinen-Wilenius, P., Tuominen, P., and Mittelmark, M. B. (2005). The development of quality criteria for research: a Finnish approach. *Health Promot. Int.* 20, 306–315. doi:10.1093/heapro/dai008

Langfeldt, L., Nedeva, M., Sörlin, S., and Thomas, D. A. (2020). Co-existing notions of research quality: a framework to study context-specific understandings of good research. *Minerva.* 58, 115–137. doi:10.1007/s11024-019-09385-2

Larivière, V., Pontille, D., and Sugimoto, C. (2020). Investigating the division of scientific labour using the Contributor Roles Taxonomy (CRediT). *Quant. Sci. Studies* doi:10.1162/qss_a_00097

Leydesdorff, L., and Wagner, C. S. (2008). International collaboration in science and the formation of a core group. *J. Informet.* 2 (4), 317–325. doi:10.1016/j.joi.2008.07.003

Martin, B. R. (1996). The use of multiple indicators in the assessment of basic research. *Scientometrics* 36, 343–362. doi:10.1007/BF02129599

Martin, B. R., and Irvine, J. (1983). Assessing basic research: some partial indicators of scientific progress in radio astronomy. *Res. Pol.* 12, 61–90. doi:10.1016/0048-7333(83)90005-7

May, R. M. (1997). The scientific wealth of nations. *Science* 275, 793–796.

Miranda, R., and Garcia-Carpintero, E. (2018). Overcitation and overrepresentation of review papers in the most cited papers. *J. Informet.* 12 (4), 1015–1030. doi:10.1016/j.joi.2018.08.006

Moed, H. F., Burger, W. J. M., Frankfort, J. G., and Van Raan, A. F. J. (1985a). A comparative study of bibliometric part performance analysis and peer judgment. *Scientometrics* 8, 149–159. doi:10.1007/BF02016933

Moed, H. F., Burger, W., Frankfort, J., and Van Raan, A. (1985b). The application of bibliometric indicators: important field- and time-dependent factors to be considered. *Scientometrics* 8, 177–203. doi:10.1007/BF02016935

Moed, H. F. (2005). *Citation analysis in research evaluation*. Dordrecht, Netherlands: Springer.

Moed, H. F. (2020). How evaluative informetrics relates to scientific, socio-historical, political, ethical and personal values. *Scholarly Assess. Rep.* 2 (1), 9. doi:10.2904/sar.18

Moher, D., Naudet, F., Cristea, I. A., Miedema, F., Ioannidis, J. P. A., and Goodman, S. N. (2018). Assessing scientists for hiring, promotion, and tenure. *PLoS Biol.* 16 (3), e2004089. doi:10.1371/journal.pbio.2004089

Moore, S., Neylon, C., Eve, M. P., O'Donnell, D. P., and Pattinson, D. (2017). "Excellence R Us": university research and the fetishization of excellence. *Palgrave Commun.* 3, 16105. doi:10.1057/palcomms.2016.105

Narin, F. (1976). *Evaluative bibliometrics: the use of publication and citation analysis in the evaluation of scientific activity*. Cherry Hill, New Jersey: Computer Horizons Inc.

National Institutes of Health (2008). NIH announces updated criteria for evaluating research grant applications. Available at: http://www.nih.gov.

Nature (2018). Science needs to redefine excellence: the concept of research excellence is ubiquitous, but its meaning depends on context (editorial). *Nature* 554, 403–404. doi:10.1038/d41586-018-02183-y

PBRF (2020). The report of the PBRF review panel; towards the tertiary research excellence evaluation. Available at: https://www.education.govt.nz/assets/Documents/Further-education/PBRF-Review/The-Report-of-the-PBRF-Review-panel-E-koekoe-te-tuie-ketekete-te-kaka-pdf (Accessed September 17, 2020).

Pendlebury, D. A. (2009). The use and misuse of journal metrics and other citation indicators. *Arch. Immunol. Ther. Exp.* 57, 1. doi:10.1007/s00005-009-0008-y

Pendlebury, D. A. (2020). "When the data don't mean what they say: Japan's comparative underperformance in citation impact," in *Evaluative Informetrics: the art of metrics-based research assessment*. Editors C. Daraio and W. Glänzel (Cham, Switzerland: Springer Nature), 115–143.

Persson, O., Glanzel, W., and Danell, R. (2004). Inflationary bibliometric values: the role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics* 60, 421–432. doi:10.1023/B:SCIE.0000034384.35498.7d

Petrovich, E. (2020). Science mapping. Available at: www.isko.org/cyclo/science_mapping.

Potter, R., Szomszor, M., and Adams, J. (2020). Interpreting CNCIs on a country-scale: the effect of domestic and international collaboration type. *J. Informet.* 14 (4), 101075. doi:10.1016/j.joi.2020.101075

Price, D. D. S. (1965). Networks of scientific papers. *Science* 149, 510–515. doi:10.1126/science.149.3683.510

REF (2020). Ref 2020/02: guidance on revisions to REF2021. Available at: https://www.ref.ac.uk/media/1417/guidance-on-revisions-to-ref-2021-final.pdf (Accessed September 17, 2020).

Roberts, G. (2003). Review of research assessment, report by sir Gareth Roberts to the UK funding bodies. Available at: http://www.ra-review.ac.uk/reports/roberts.asp (Accessed September 17, 2020).

Rogers, G., Szomszor, M., and Adams, J. (2020). Sample size in bibliometrics. *Scientometrics.* 125, 777–794. doi:10.1007/s11192-020-03647-7

Schubert, A., and Braun, T. (1996). Cross-field normalization of scientometric indicators. *Scientometrics* 36, 311–324. doi:10.1007/BF02129597

Schubert, A., and Braun, T. (1993). Reference standards for citation based assessments. *Scientometrics.* 26, 21–35. doi:10.1007/BF02016790

Sivertsen, G. (2016). "Publication-based funding: the Norwegian model," in *Research assessment in the humanities: towards criteria and procedures*. Editors M. Ochsner, S. E. Hug, and H. D. Daniel (Zürich, Switzerland: Springer Open), 79–90.

Sivertsen, G., Rousseau, R., and Zhang, L. (2019). Measuring scientific contributions with modified fractional counting. *J. Informet.* 13 (2), 679–694. doi:10.1016/j.joi.2019.03.010

Sivertsen, G. (2018). The Norwegian model in Norway. *J. Data Infor. Sci.* 3 (4), 3–19. doi:10.2478/jdis-2018-0017

Sivertsen, G. (2017). Unique, but still best practice? The research excellence framework (REF) from an international perspective. *Palgrave Commun.* 78, 1–6. doi:10.1057/palcomms.2017.78

Small, H. (1973). Co-Citation in the scientific literature: a new measure of the relationship between two documents. *J. Am. Soc. Inf. Sci.* 24, 265–269. doi:10.1002/asi.4630240406

Small, H., and Griffith, B. C. (1974). Structure of scientific literatures. I: identifying and graphing specialties. *Sci. Stud.* 4, 17–40. doi:10.1177/030631277400400102

Tang, L., Shapira, P., and Youtie, J. (2015). Is there a clubbing effect underlying Chinese research citation increases? *J. Assoc. Infor. Sci. Tech.* 66 (9), 1923–1932. doi:10.1002/asi.23302

Thomas, D.A., Nedeva, M., Tirado, M.M. and Jacob, M. (2020). Changing research on research evaluation: A critical literature review to revisit the agenda. *Res. Eval.* 29(3), 275–288. doi:10.1093/reseval/rvaa008

UGC (1984). *A strategy for higher education into the 1990s*. London, United Kingdom: H.M. Stationery Office.

UGC (1982). Report of a Joint working party on the support of university scientific research (the Merrison report). Cmnd 8567.

Van Eck, N. J., and Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* 84, 523–538. doi:10.1007/s11192-009-0146-3

Van Leeuwen, T. N., Moed, H. F., Tijssen, R. J. W., Visser, M. S., and van Raan, A. F. J. (2000). First evidence of serious language-bias in the use of citation analysis for the evaluation of national science systems. *Res. Eval.* 9, 155–156. doi:10.3152/147154400781777359

Van Leeuwen, T. N., Moed, H. F., Tijssen, R. J. W., Visser, M. S., and van Raan, A. F. J. (2001). Language bias in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance. *Scientometrics.* 51, 335–346. doi:10.1023/A:1010549719484

Wagner, C. S., and Leydesdorff, L. (2005). Network structure, self-organization, and the growth of international collaboration in science. *Res. Pol.* 34, 1608–1618. doi:10.1016/j.respol.2005.08.002

Wagner, C. S. (2008). *The new invisible college*. Washington, DC, United States: Brookings Press.

Waltman, L. (2016). A review of the literature on citation impact indicators. *J. Informet.* 10, 365–391. doi:10.1016/j.joi.2016.02.007

Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E. C. M., Tijssen, R. J. W., and van Eck, N. J. (2012). The Leiden ranking 2011/2012: data collection, indicators, and interpretation. *J. Am. Soc. Inf. Sci. Technol.* 63 (12), 2419–2432. doi:10.1002/asi.22708

Waltman, L., and van Eck, N. J. (2015). Field-normalized citation impact indicators and the choice of an appropriate counting method. *J. Informet.* 9 (4), 872–894. doi:10.1016/j.joi.2015.08.001

Waltman, L., and van Eck, N. J. (2016). "The need for contextualized scientometric analysis: an opinion paper," in *Proceedings of the 21st international conference on science and technology indicators*. Editors I. Ràfols, J. Molas-Gallart, E. Castro-Martínez, and R. Woolley (València, Spain: Universitat Politècnica de València), 541–549.

Weinberg, A. M. (1963). Criteria for scientific choice. *Minerva.* 1, 159–171. doi:10.1007/BF01096248

Wilsdon, J., Allen, E., Belfiore, E., Campbell, P., Curry, S., and Hill, S. R., Kain, R., Kerridge, S., Thelwall, M., Tinkler, J., Viney, I. and Wouters, P. (2015). The metric tide: independent review of the role of metrics in research assessment and management. London: SAGE Publications Ltd. (Accessed January 7, 2021).

Zitt, M., Ramanana-Rahary, S., and Bassecoulard, E. (2005). Relativity of citation performance and excellence measures: from cross-field to cross-scale effects of field-normalization. *Scientometrics* 63, 373–401. doi:10.1007/s11192-005-0218-y

Conflict of Interest: The authors are members of the Institute for Scientific Information, which is part of Clarivate, a company that also owns the Web of Science which is the data source used throughout this article.

frontiers
in Research Metrics and Analytics

ORIGINAL RESEARCH
published: 19 April 2021
doi: 10.3389/frma.2021.643533

Dimensional Taxonomy of Data Visualization: A Proposal From Communication Sciences Tackling Complexity

Victor Cavaller *

Department of Information and Communication Sciences, Universitat Oberta de Catalunya (UOC), Barcelona, Spain

This article consists of a conceptual analysis—from the perspective of communication sciences—of the relevant aspects that should be considered during operational steps in data visualization. The analysis is performed taking as a reference the components that integrate the communication framework theory—the message, the form, the encoder, the context, the channel, and the decoder—which correspond to six elements in the context of data visualization: content, graphic representation, encoding setup, graphic design and approach, media, and user. The study is focused accordingly on the dimensions that these elements describe: the degrees of abstraction of the information, the functionalities of the tool for the graphical representation, the specifications for the setup of the visualization, the approach modes to the context by the graphic design, the levels of communication efficiency in the media, and the requirements of the visualization perceived as values from the user experience side. The unfolding of these dimensions is undertaken following a common pattern of six organizational layers of complexity—basic, extended, synthetic, dynamic, interactive, and integrative—according to the analytical criteria. The results of the detailed study, based on an extensive scientific literature review, allow the design of a dimensional taxonomy of data visualization built on a matrix structure where these elements act as factors of completeness and the layers act as factors of complexity. As a conclusion, an object-centered model constituted by an ordered series of phases and achievements is proposed as a guide to complete a systematic process of data visualization.

Keywords: index terms: data visualization, dimensional taxonomy, communication process, communication theory, knowledge transfer, complexity

OPEN ACCESS

Edited by:
Juan Ignacio Gorraiz, Universität Wien, Austria

Reviewed by:
Chao Min, Nanjing University, China
Francis Harvey, Leibniz Institute for Regional Geography (IfL), Germany

Correspondence:
Victor Cavaller
vcavaller@uoc.edu

Specialty section:
This article was submitted to Research Assessment, a section of the journal Frontiers in Research Metrics and Analytics

Received: *18 December 2020*
Accepted: *09 February 2021*
Published: *19 April 2021*

Citation:
Cavaller V (2021) Dimensional Taxonomy of Data Visualization: A Proposal From Communication Sciences Tackling Complexity. Front. Res. Metr. Anal. 6:643533. doi: 10.3389/frma.2021.643533

INTRODUCTION

Complexity as a Challenging Parameter to Integrate in Data Visualization

Over the past decades, visualization and complexity have received extensive scientific attention, and there has been a huge increase in the number of publications dealing directly or indirectly with their relation. Emergent complexity in systems theory is described as the distinctive novel properties or behaviors that arise in organizations from the interaction among their components (Gibb et al., 2019). Adding complexity is the common response of organizations under the influence of controllable and uncontrollable factors, by means of which they adapt themselves to changes in the environment. In complex systems, emergent properties are provided by networks of internal processes and hyper-processes in order to accomplish a particular function, which means that there is a scale factor involved in their structure. Complexity is deeply embedded in organizational dynamics, and it has become a real challenge for data visualization. If complexity characterizes in general any organization or phenomenon, by extension, the methods and techniques to visualize them must be accordingly modified or eventually adapted to capture the dimensional structure and scaled dynamics that configure the object.

Among the fields in which publications about complexity have reached more popularity in the last few years, modeling of biological ecosystems (May, 2019), social complexity (DeLanda, 2019), self-organization in statistical mechanics (Wolfram, 2018), ecological complexity (Allen and Starr, 2017), and economic complexity (Hausmann et al., 2014) can be highlighted. In a similar review focused on visualization, publications in cartography (Kraak and Ormeling, 2020), perception and design (Ware, 2019), sequencing technologies (Katoh et al., 2019), molecular visualization and analysis (Goddar et al., 2018), flow visualization (Yang, 2018), multivariate density estimation (Scott, 2015), neural networks (Yosinski et al., 2015), genetics (Hu et al., 2014), information and knowledge (Börner, 2015), and system visualization, analysis, and design (Munzner, 2014) have obtained an overwhelming impact.

In the particular field of data visualization, from its very beginning, pioneering works from authors such as Bertin (1967), Tufte (1983), Schneiderman (1996), Horn (1998), and Wilkinson (1999), followed by precursors such as Fayyad et al. (1996), Hoffman et al. (1997), Hoffman et al. (1999), Fayyad and Uthurusamy (2002), Manovich (2001, 2011, 2013, 2016), and Few (2011) up to the most prominent authors in the field such as Cairo (2013), Keim et al. (2013), Börner and Polley (2014), Heer and Agrawala (2006), Heer and Scheiderman (2012), Viegas and Wattenberg (2006), Wattenberg et al. (2016), or Benoit (2019) with many others, have made a great effort to ground data visualization on scientific principles.

The Lack of an Integral Data Visualization Taxonomy to Tackle Complexity

Data visualization and complexity as scientific topics are undergoing a period of consolidation with an increasing and overwhelming number of scientific publications and specialists working on these fields. However, along with this positive impression, a more detailed overview suggests that linked problems remain unsolved:

1) From the object side, at any scientific discipline where the concept of complexity appears, it refers to objects constituted by interconnected layered networks; however, there is not a common proposal for a pattern of complexity in phenomena from both an organizational and analytical perspective.
2) From the subject side, the current limits dealing with the issue of complexity do not lie so much in its evidence or even characterization but rather have to be sought in the fact that the complexity of the object obliges the particular adaptation and sophistication of data visualization methods which call for a definition of their analytical potential to describe it.
3) Finally, from the practice side, data visualization is driven by highly demanding standards—its universal application as a tool, its specialization and versatility, and its need for effective and immediate results.

The root cause of the above-mentioned problems is the absence of an operative standard for the implementation of data visualization. As a consequence, the main deficit, repeatedly observed throughout this review, is that data visualization is still affected by a serious lack of systematicity which ultimately—from the perspective of communication sciences—can be summarized as the lack of an integral taxonomy.

There is no science without its own taxonomy. Taxonomy is the practice used by any science to clarify itself by classifying its concepts, being thus an exercise of self-explanation about its fundamentals. Data visualization occupies a central position as an applied science—in an intersection among statistics, semiotics, computer science, graphical design, and psychology, in close relation to communication sciences—which means that the meta-analysis required in order to generate a taxonomy must be performed over multiple scientific disciplines. Being central paradoxically represents a weakness.

Despite the fact that there have been tentative approaches to define a taxonomy in particular areas of data visualization (Schneiderman, 1996; Heer and Shneiderman, 2012; Ruys, 2020), the critical requirement for an integral taxonomy is a pending workload, and it is currently having a negative impact on both its consolidation as a rigorous technical method and on its recognition as a scientific discipline, beyond its instrumental use. Faced with this situation, it is appropriate to shed light on the foundations of the discipline of data visualization—understood as a communication process—in order to provide a solid ground for its systematic application. To achieve such purpose, a key action is required. Complexity has to be integrated as an internal parameter in the configuration of its operative. As complexity is a factor that constitutes the object and conditions the subject, data visualization needs to undergo a conceptual analysis object-centered on organizational complexity, which in turn must be tracked to each of the components of the communication process that participates in data visualization. This article is focused on this objective.

Communication Components and Layers of Complexity in the Data Visualization Process

Any scientific research inquiry follows three procedural stages when managing data: data formalization, data analysis, and data visualization, which, respectively, transform observations and measurements into data, data into information, and information into knowledge. Formal data appear as a result of preprocessing operations, information appears as a result of data analysis, and knowledge appears as a result of data visualization.

Data visualization can be transversely used as a tool in both processes of data formalization and data analysis, but ultimately, it constitutes the final and synthetic visible stage where the results of data analysis are reported. In fact, by means of the accuracy of data visualization, the success of any data processing is evaluated. In order to provide instruments from communication sciences that can contribute to the process of transforming data into understandable information and information into valid knowledge, it is necessary to deal with data visualization in a systematic way covering the totality of the factors that are involved in its process.

The first step to start a thorough review of these factors is to identify the following elements that participate in data visualization understood as a communication process:

- the *content*, the data, and information to be communicated
- the *graphic representation* of this content
- the *encoding* of the information integrating data and graph specifications
- the *design* adapted to the *context*, the audience, or the target
- the *media* by which the visualization is published and disseminated
- the *user* who receives the visualization

The proposal of these elements is not arbitrary. "Data visualization uses principles, concepts, techniques, and theories that come from multiple backgrounds: programming, web design, semiotic, or psychology" (Aparicio and Costa, 2014). However, from the point of view of the communication theory, these core elements are embedded in data visualization, beyond its background and application, in so far as they correspond to the most widely accepted framework of the communication model (Shannon and Weaver, 1949; Schram, 1954; Berlo, 1960; Rothwell, 2004; Barnlund, 2008). The elements are as follows:

1) MESSAGE: *what* things are communicated according to the context
2) FORM: *how*, in which form and by which *tool*, the content is communicated, taking into consideration the media through which it is broadcasted
3) SENDER: *by whom* and *why* the communication is provided and encoded in a singular setup observing the receiver
4) CONTEXT: *what* is the scenario *where* the communication takes place
5) CHANNEL: *through which* medium is the information communicated
6) RECEIVER: *to whom* the information is addressed

These six elements must be considered as factors of completeness in data visualization. The failure to observe any of them is a recurring cause of miscommunication and misunderstanding. Data visualization constitutes a process of communication, the efficiency of which is conditioned by the actions that these elements imply: the selection of the content, the formal representation of the information, the encoding and setup of the visualization, the graphical design appropriate to the context, the adaptation to the medium, and the observation of user preferences. Furthermore, understanding the completion of these actions as a critical success factor, they must be undertaken considering their interconnection which plays a critical role and can be expressed by means of the following practical questions:

1) What content does the sender want to communicate and to what *degree* of abstraction?
2) In which form? Which *functionalities* from which tools are appropriate for the graphical representation to be integrated in the pursued channel?
3) Once content and form are defined, what *specifications* must be applied to the setup of both data and graphical representation in order to adapt to each other?
4) What are the approach mode and the graphical design suitable to the context? What *properties* does the visualization have to meet depending on the target or audience?
5) What characteristics must the visualization contemplate in order to make it efficient according to the media where it is projected? What are the *levels* of communication efficiency that must be achieved?
6) What *requirements* must be observed from the user's experience in order to improve understanding of the topic?

Objectives and Method: Building a Taxonomy

The above questions highlight six *dimensions* of the communication process that, conditioning the systematic procedure of data visualization, must be accurately studied:

- the degrees of abstraction of the information
- the functionalities of the tool for the graphical representation
- the specifications for the setup of the visualization
- the approach modes to the context by an appropriate graphic design
- the levels of communication efficiency in the media
- the requirements of the visualization perceived as values from the user experience side

The definition of these dimensions leads to the equally important issue of internal order in which they must be unfolded. From previous studies about data analytical procedure (Cavaller, 2008; Cavaller, 2007), it has been shown that, as a general rule, the construction of indicators applied to data analysis is correlated with the layers of organizational complexity that exist in any organized entity or phenomenon:

1) Basic layer: basic interactions
2) Extended layer: multivariate relationships
3) Dynamic layer: distributions or multi-relational dynamic
4) Synthetic layer: internal logics or processes
5) Interactive layer: system as architecture of hyper-processes
6) Integrative layer: organization as ecosystem

TABLE 1 | Matrix architecture of factors of completeness and complexity for the design of the dimensional taxonomy of data visualization according to the components of the communication framework theory.

		Communication framework component					
	Elements	**Message**	**Form**	**Encoder**	**Context**	**Channel**	**Decoder**
	Categories	What	How	By whom &why	Where and when	Through which	To whom
				Factors of completeness			
	Elements	**1** **Content**	**2** **Graphic Representation**	**3** **Encoding Set-up**	**4** **Graphic Design and approach**	**5** **Media**	**6** **User**
	Dimensions	**Degrees** of abstraction	**Functionalities** of the tools for the graphical representation	**Specifications** of the set-up of the visualization	Approach **modes** and **properties** of visualization	**Levels** of communication efficiency	**Requirements** from the user experience side
	Layers			**Questions for the dimensional analysis**			
Factors of complexity	1 **Basic** 2 **Extended** 3 **Dynamic** 4 **Synthetic** 5 **Interactive** 6 **Integrative**	What content do you want to communicate and to what degree of abstraction?	In which form? Which functionalities from which tools are appropriate for the graphical representation that is pursued?	What specifications must be applied to the setup of both data and graphical representation in order to adapt each other?	What are the approach modes and the graphical design suitable to the context, target or audience?	What characteristics must be contemplated to achieve the levels of communication efficiency according to the media?	What requirements must be observed from the user's experience in order to improve understanding of the topic?

Given that the layers of complexity of any object or phenomenon condition the structure of the analytical procedure, data analysis imposes a scale approach on data visualization in an object-centered way. Consequently, the sequential and detailed unfolding of data visualization—covering degrees, functionalities, specifications, modes and properties, levels, and requirements—must be internally described through cross-cutting layers.

Taking this conception as a starting point of the review and the analysis, the goal of this article was to design an *object-centered data visualization model*, organized in two axes:

- as a set of gradual approaches to the complexity of the dimension that is managed
- by means of the progressive completion of the corresponding communication component

As a result, a *dimensional taxonomy* of data visualization based on a matrix structure—where the elements that participate in data visualization act as factors of completeness, and their development in layered dimensions act as factors of complexity—is proposed (see **Table 1**).

It must be observed that building the proposed taxonomy, the *theoretical framework* of communication sciences is projected as the *practical framework* for the dimensional analysis of data visualization. Meaning that in order to validate it, this article has been focused on an extensive systematic review of the scientific literature and on a conceptual analysis about the relevant aspects that have been considered both in practice and in the current debates about data visualization, categorizing them into topical groups taking as a reference those components and layers.

CONTENT AND DEGREES OF ABSTRACTION OF INFORMATION

The first node of the communication framework is the message or the content of the communication. The first of the main functions of data visualization is to communicate a message: generally, information about an event, a phenomenon, a process, a system, or in general, any observable subset of the real world. At this starting stage, the assumption of the quality of data about the object is accepted as a fact because it should result from previous tasks of data formalization and analysis. Data visualization, from the perspective of the content to be represented, must distinguish six degrees of abstraction of information which correspond to six layers of organizational complexity.

Parameters, Sample, and Descriptive Statistics

In practical terms, data visualization can be faced with three potential initial scenarios: a requirement of data visualization without previous data formalization, without previous data analysis, or, in the best case, with both data formalization and analysis previously performed. In the first scenario—that could be called agile, *adhoc*, or express demand—data visualization procedure must introduce a delay to examine the target in detail, to seek evidence, and to detect the different properties which presumably can be sustained by available data, in order to complete a proper answer to the requirement. The so-called data wrangling or data preprocessing operations are required before data analysis; such operations include data cleaning, matching, organization, and aggregation (Chen et al., 2015). In the second scenario, once a formalized dataset has been obtained or is available from a system of information, the actions to be carried out can directly jump to check whether the target can be delimited and whether a reduced and representative sample for a deeper analysis is available. In the third scenario, as the attention has already been focused on the particular issue, the consequent step is to select the data and constitutive relations that adequately answer the visualization requirement. In any case, evidence must exist and must be reducible to parameters and measurable. The *congruence* as the essential quality of being in agreement with the real-observed facts should be the principal and basic characteristic of data visualization.

Clustering of Parameters: The Construction of Indicators as Evidenced Relations

A second degree of abstraction of information is reached when the requirement for data visualization needs adding and accumulating new observed properties about the subject to the focus. Different dimensions, traits, or aspects about the same reality are defined by different parameters as variables in a way that aggregating them by mathematical calculations leads to the construction of indicators that make their relations visible which "you otherwise would have been blind to if you looked only at the naked source" (Yau, 2013). The process of aggregating variables describing parametrical relations needs a thorough investigation, comprehensive in scope. A formal condition of this clustering can be defined as *exhaustivity*, the need to address all aspects without omission.

Multi-Relational Dynamics: Set of Variables' Distribution

The next degree of abstraction of the information is focused on the dynamics which refers to the multiple and observable distributions and relationships between sets of variables. It is understood that prior to data visualization, data analysis has been carried out in terms of detecting correlation or causality between variables. The definition of the relationships, as patterns in the dynamics, between sets of variables is considered as explanation of the variations observed in the phenomenon. A pattern is defined as any regularly repeated arrangement or relation in or between a set of parameters that modifies others or changes itself according to its distribution. Among all reasonable explanations, the best one covers the greatest spectrum of observed relationships or fits well enough to a sufficient portion of all the available information. The *consistency* is the modal quality—of being in the harmony, compatibility, and uniformity—that the explanation with the observation of particular distributions should pursue when dealing with the content of data visualization.

Conceptual Synthesis and Symbolic Abstraction of a Process

In case of wanting to visualize a complex phenomenon, usually associated to a process, the definition of the parameters, the construction of indicators, or the detection of interconnected factors or patterns is not enough because the abstraction required is an, more than probable, explanation. Explaining a phenomenon as a set of separated dynamics is not sufficient either. The fourth degree of information abstraction involves the conceptualization of the internal relationship, the sequential process, and the vector direction that describes a phenomenon or lies behind the events. The nature of the interconnection between the dimensions of a process has to be observed as an objective condition of having a *logical unity in coherence*. When an explanatory model is involved as a communication message, data visualization requires a previous conceptualization, summarizing the accepted premises about the object logically interconnected.

Layered Processes, Hyper-Processes, and Systems: Experimentation and Testing Hypothesis

When considering systems where hyper-processes—resulting from the coexistence of interconnected processes—are involved, a higher degree of abstraction in information must be achieved. The internal complexity of a phenomenon needs the definition of the layers where each constituent process takes place. The object of data visualization at this level goes from what was initially perceived as an isolated process to its interaction with other processes that condition each other, defining a network of system functions and their interactions.

Figure 1 shows the graphical representation of data on the layers of parallel activities undertaken by a university, illustrating how they participate in scientific research and technological development. The multilayered structure describing a hyper-process model is a clear expression of the crucial ability of systems to adapt to the complexity of the changing environment.

Scientific progress implies the proposal of competing explanatory models, the certainty of which cannot be achieved. So there being no *verifiability* but *falsifiability* by experimentation (Popper, 1959), the evaluation of the confirmatory or falsifying value of evidence about a hypothesis depends on their demonstrative condition, which data visualization must facilitate in order to achieve scientific consensus.

Abstraction in Scientific Modeling as a Reconstruction of an Organization

The degree of abstraction of the information is correlated with the complexity of the entity from which data have been obtained and data visualization has to show. The procedure of grouping a network of interactive processes in different layers is definitely dealing with the highest level of complexity that culminates the scope of data visualization in which an organization within its environment is explained. Scientific *modeling* and simulation are the results of a simplification and abstraction of human perception and conceptualization of reality that in turn come from physical and cognitive constraints. Modeling allows scientists to implement their *reconstruction*, simulating the program or code of the organization, future behaviors, visualizing scenarios, manipulating, and gaining intuition about the entities, phenomena, or processes being represented, for managerial or technical decision-making. At this level, uncertainty is a transcendent condition characterized by limited knowledge which ranges just beyond the experimentation in order to achieve a holistic view of a phenomenon.

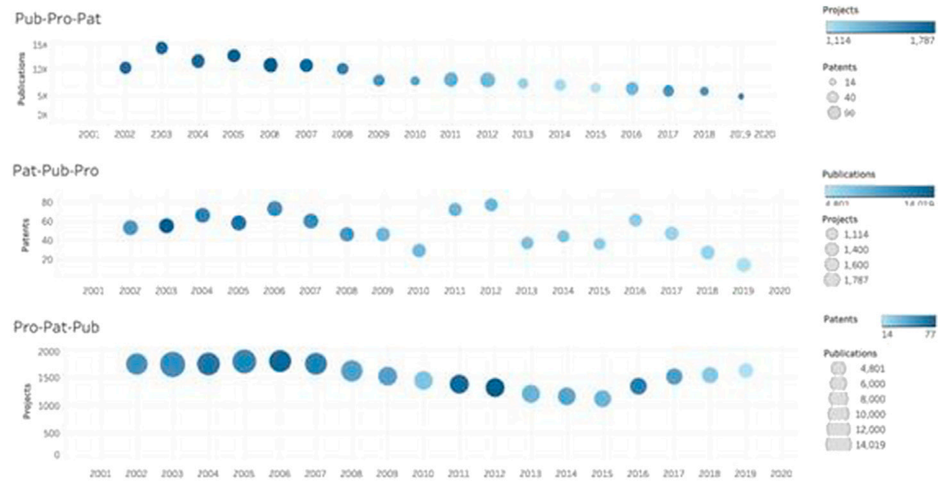

FIGURE 1 | Detail of the interactive diagram map of the evolution of projects, patents and publications, considered as layers of parallel outcome processes of UPC in the period 2002–2019. Source: Cavaller et al. (2020), Own elaboration based on Future UPC data with Tableau.

GRAPHICAL REPRESENTATION AND TOOL FUNCTIONALITIES

Visualization has been defined as "a transformation of quantified data which is not visual into a visual representation"—"a remapping from other codes to a visual code" (Manovich, 2001; Manovich, 2011). Once the answer to the ominous question—which data in which degree of abstraction related to which level of organizational complexity about which object is required to be represented—is clear, the next question is: What is the ideal graphic representation to visually transform these data with a strictly functional orientation? This decision is not trivial. Principles of graphic communication, studied by semiology or semiotics, under which diagrams, networks, and maps or any sign in general are used, have been designed for the production of meaning in their close relation to the analysis of the information that they represent (Bertin, 1967). Here, it is worth remembering that one of the most recurrent errors in data visualization is to confuse the criteria for the selection of a proper graphic representation of data with the criteria of the graphic design of the visualization. The graphic representation from a functional point of view is directly related to the nature of the content to be displayed. In this sense, six different *object-oriented* graphic representations with six different functionalities can be defined.

Basic Functionalities for a Descriptive Graphical Representation

The basic functionalities in the graphical representation of data are associated with a descriptive visualization of the parameters that depict a phenomenon. Information visualization relies on two key principles: reduction by the use of graphical primitives such as points, straight lines, curves, and spatial variables such as position, size, shape, or movement "to represent key differences in the data and reveal patterns and relations," privileging them over other visual dimensions (Manovich, 2011). The development and formalization of statistical tools for the analysis and graphical representation of data have had a great impact in the field of visualization (Friendly, 2006). Graphs are useful to show relationships among variables—how a whole is divided into different parts, how variables have changed over time and their range, when and how data are connected, what are the trends, and how changes in one variable affect another—or to obtain a sequence in the development and transformation of trends or patterns. The main quality that is required from a graphic representation is to be descriptive. In this sense, the *evidentiality*, the condition to provide evidence, in an illustrative, expressive, and depictive way, is an essential condition by means of which the quality of the graphical representation in data visualization is evaluated.

Advanced Functionalities for a Relational Graphical Representation

Multivariate or relational visualization involves the observation of multiple measurements and their relationship. There are different methods of visualizing a multidimensional or multivariate reality capable of covering a wide spectrum of inputs and outputs, associated with different analysis techniques and methodologies. "Data can be aggregated in many ways before being visualized in charts, profoundly affecting what a chart conveys" (Kim et al., 2019). In general, the need to express comparison, correlation, distribution, proportions, and hierarchy relationships in a dataset requires advanced functionalities in the visualization design. Two of the main principles of graphical integrity defined by Edward Tufte are referred to as *proportionality* and disambiguitty. "The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities represented. Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity" (Tufte, 1983). The property that pursues advanced forms of graphic representation is *integrity*, the formal condition of maintaining a direct proportion in the scale relationship of the

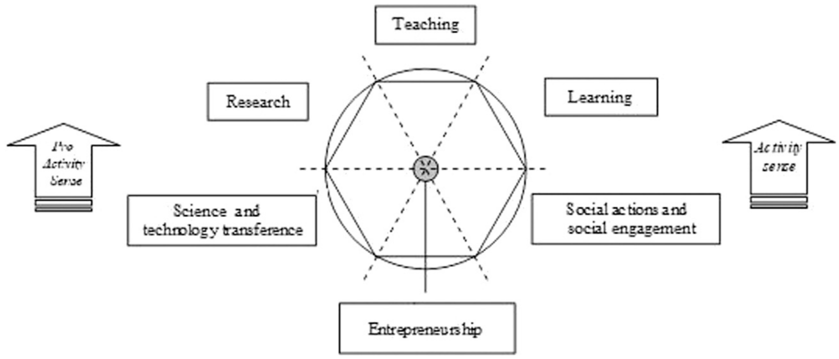

FIGURE 2 | The Sextuple Helix Model of the KT sequential and cyclical activities in higher education (Cavaller, 2020).

parts with the whole and with the unit of measurement, without distorting the degree of interdependence of the variables.

Functionalities for a Graphical Representation in a Dynamic Visualization

Dynamic or multi-relational visualization represents a reality where all the factors—defined as set of related parameters—are interconnected, and therefore there is interdependence between them, and consequently, their network position changes according to a spatial or temporal joint distribution. "Data visualization is an efficient means to represent distributions and structures of datasets and reveal hidden patterns in the data" (Chen et al., 2015). A dynamic visualization of data has to be facilitated by functionalities of tools for the transmission and understanding of the global and interconnected networked nature of a reality that is itself dynamic. Modeling of dynamic interaction networks has traditionally been supported by graph stream techniques or dynamic graph models (Harary, 1969). Among multiple applications, data visualization is useful for social learning analytics providing "additional information about actors and their behaviors for decision-making in online distance learning" (Hernández-Garcia et al., 2015) The *schematicity*, the ability to present in a schematic way the main features of the connectivity of the phenomena, allows the evaluation of the dynamic graph quality.

Process Graph, Info Graphics, and Motion Graphics: Representing Processes

Process visualization must describe the internal logics that lie behind phenomena. Once the interdependence relationships between the different factors or dimensions of a phenomenon are known, the existence of its internal logic can be inferred, and therefore it is possible to define an explanatory model and proceed to its visualization. However, in order to obtain a synthetic visualization that brings together the different dynamic perspectives of the same reality, continuing with a quantitative gradation in the abstraction of information and with its corresponding visualization is meaningless or clearly insufficient. The natural path to parameterization and visualization requires a qualitative leap that is made through symbolic abstraction with the use of info graphic representation and animation techniques (Harrison et al., 2015). The graphic representation at this level of complexity is done through process graphs, graph processing workflows, info graphics, and motion graphics (Curcin et al., 2010; Riazi and Norris, 2016; Microsoft Visio, 2020), or by means of diagrams, maps, or info graphics, it can display chronological, comparative, flow diagrams, anatomical, statistical, geographical, hierarchical or hybrid forms (Cairo, 2013; Manovich, 2016; Inequaligram, 2020). The quality of the process of graphical representation is defined by its objective condition of expressing the logic of their transformation flow, which refers to its *sequential* or *flow logicality*. In **Figure 2**, a process graph describing the Sextuple Helix Model for the assessment of universities based on KT processes is shown. The activities—as nodes—are proposed, in a two-way cyclical sequence, for their correspondent accounting and mission values (Cavaller, 2020).

Interactive Graphics: Hyper-Process and System Graphical Representation

What is the ideal form of graphical representation of a phenomenon when its internal logic is also changing, complex or/and simultaneous processes in different layers at different scales are interrelated? Hyper-process or system graphical representation is needed to observe the constituent layers when describing the architecture of the systems where different processes coexist. The graphic representation of the data at this level must allow the user to interact with the visualization in order to know independently or together the different layers that are integrated into the phenomenon and their connection. One of the most common forms of graphic representation for interactive visualization is the interactive map. In general, interactive graphics are a type of graphic representation, which points to a demonstrative condition of an explanatory model, the ability to show evidence that verifies or refutes a hypothesis or theory that is defended. This property of *complex evidentiality* referred to a graphical representation means that its quality is evaluated by a demonstrative condition, the ability to give detailed, interactive, and *ad hoc* access to evidence of complexity, in order to demonstrate all the factors of a theory that is defended.

Convergence of the Symbolic and Analytical Path: Integration of Scientific Data Visualization

Multidimensional phenomena structured in different layers of processes, where different organizational systems are involved, make their graphic representation an extremely complex matter. This difficulty has led to the need to develop new functionalities in the visualization tools that allow a comprehensive and holistic representation. At this level of complexity, the convergence between the symbolic and the analytical paths in data visualization prevails and requires that the graphic representation of data be accompanied in turn by a figurative or symbolic visual reconstruction of the reality of the phenomenon. This requirement is easily observable in scientific visualization, such as in modeling projects of biological systems (Ambicon, 2016) or in the field of medicine (Jang et al., 2014), to name a few examples, thanks to the extensive use of computer aided systems (SystemsBiology, 2016). The quality of a scientific visualization is evaluated by the capacity of reconstitution of reality in its entirety, even in what is not known in detail. The graphic representation associated with scientific modeling and simulation is characterized by the ability to represent the *intricacy* of phenomena internally and in their relationship with their environment, reconstructing those elements in which no sufficient evidence is available and posing them for a future demonstration.

ENCODING SPECIFICATIONS AND CONFIGURATION SETTINGS

The third node of the communication framework is the encoder, and its action is the encoding or the communication configuration. The main function of data visualization associated with this node is to communicate, so as to add meaning to the data and transform it into information. Visualization is the discipline that consists of "transforming data into meaningful information" (Benoît, 2019), and this transformation is made by encoding.

Data Formalization *Adhoc*. Setting up Data and Plotting Elements for Descriptive Visualization

The first step in the basic configuration of data visualization is to specify and verify basic elements selected such as parameters, constants and variables, scale, data range, sample, legend, and labels, and to check them in preliminary views in order to manipulate and ensure the accuracy of the representation. The basic operations to be performed in this phase have been proposed—in a synthetic way—as tasks grouped into three high-level categories: 1) specification of data and views (visualize, filter, order, and derive), 2) view manipulation (select, navigate, coordinate, and organize), and 3) process of analysis and provenance (record, note, share, and guide) (Heer and Shneiderman, 2012). Visual analysis tools, such as Profiler, have been designed for assessing quality issues in tabular data such as missing, erroneous, extreme, and duplicate values that undermine analysis and are time-consuming, applying "data mining methods to automatically flag problematic data" and suggesting "coordinated summary visualizations for assessing the data in context" (Kandel et al., 2012). The specifications for a basic encoding of the data and its graphical representation pursue *accuracy*, an essential quality of being correct or precise for a basic visualization.

Multidimensional Transformation

The configuration for the visualization of a multivariate set is basically solved in its transformation to a data matrix with rows and columns, representing cases and variables. There are different theoretical approaches or models that describe the procedural stages of configuring data visualization. "Card's early model lists four successive steps: 1) the processing of raw data, 2) the transformation of data tables, 3) the mapping of visual structures, and 4) the transformation of the visual results (e.g., zooming and overview)" (Vande Moere and Purchase, 2011). Multidimensional transformation is related to the concept of visual metaphors and to the capacity for interaction (Kosara et al., 2006). The specifications for the configuration of multidimensional data and its graphical visualization pursue the preservation and detailed rigor of the proportions in the relationships detected between variables (Blackwell, 2011; GIS, 2020). In multidimensional data representation, such as 3D scatterplots, this is done by using a software volume renderer for display, combining it with InfoVis interaction methods such as linking and brushing (Kosara et al., 2004), selecting or displaying subsets of data and defining the relationship between them. In the Card's model, novel visual metaphors represent the structure of, and the relationships within, complex data (Card et al., 1999; Vande Moere and Purchase, 2011). The property or quality that is pursued in a multidimensional configuration of data visualization is *multidirectionality*, a formal condition, which is defined as the ability to show the widest possible range of interrelationships between set of variables.

Configuration of Data and Representation of Dynamic Multidimensional Distribution in Data Visualization: Integration of Applications

The next step in the process of configuring data visualization focuses on the dynamic relationships and distributions between groups of variables, combining and communicating different visualization techniques and methodologies, which generates a fundamental requirement for dynamic, compatible, and interconnected tools for visual encoding. "Tools do not exist in isolation, but within an ecosystem of related components" (Bostock et al., 2011). New tools have been designed to facilitate application integration in visualization design. However, "despite a diversity of software architectures supporting information visualization, it is often difficult to identify, evaluate, and reapply the design solutions implemented within such frameworks" (Heer and Agrawala, 2006). One successful example of this capability is dashboards, very useful embedded tools that allow the programmer to develop visualizations of known variables, dimensions, and relationships from a dataset (Klipfolio, 2020). Dashboards combine "multiple conventional data visualization styles to most efficiently and accurately be able to understand data," "facilitating exploratory analysis and answering a multitude of new questions" (McCann, 2020). The difficulties in detecting and making the behavior patterns of dynamic distribution visible are associated with the difficulties of integration of the visualization tools (Heer and Agrawala, 2006). The modal quality that is required from the configuration of a dynamic visualization is its *versatility* in terms of interconnectivity and compatibility with other tools.

Configuring Data Process Visualization

Following consecutive levels of complexity, the configuration is directed to the design and programming of algorithms that simulate the operation of the logical structure of the process that underlies the phenomenon to be represented graphically. The modeling of a process, in order to visualize it, includes different operational moments: 1) defining the flow diagrams and the forms of representation, 2) selecting inputs and outputs of the processes for each of the events and activities, and 3) obtaining or designing the algorithms that synthetically define their relationship in the analyzed process. The configuration has to point to the definition of an explanatory model that is represented and, therefore, to the logical structure that underlies (Curcin et al., 2010). An example of a software tool that allows the configuration of process visualizations by generating algorithmic art is processing (Reas and Fry, 2007; Terzidis, 2009; Greenberg et al., 2013; Processing, 2020). In machine industry and manufacturing methods, control systems such as the Supervisory control and data acquisition (SCADA) incorporate graphical user interface (GUI) and allow users to interact with electronic devices, computers, networked data communications through graphical icons, and audio indicator (Boyer, 2010; Siemens, 2020). The property that is sought in a configuration of the visualization of a process is being *self-explanatory*, an objective condition of being able to express the autonomous mechanics of a process easily understood.

Specifications for the Configuration of the Data and Its Graphical Representation in an Interactive Visualization

Getting into the internal complexity of phenomena involves defining the different layers of sub- or super-processes that participate or overlap in strata, which in turn requires developing and mastering complex visualization tools. The specifications for the configuration of an interactive visualization are framed in the experimental and demonstrative stages of the research. Tools such as Vega-Lite—"a high-level grammar of interactive graphics"—allow the generation of visualizations to support analysis, supporting "data transformations such as aggregation, binning, filtering, sorting, and visual transformations including stacking and faceting," composing specifications "into layered and multi-view displays, and made interactive with selections" (Vega-Lite, 2020). "Users specify interactive semantics by composing selection" abstractions that define "input event processing, points of interest, and a predicate function for inclusion testing. Selections parameterize visual encodings by serving as input data, defining scale extents, or by driving conditional logic" (Satyanarayan et al., 2017). "When building visualizations, designers often employ multiple tools simultaneously (Bostock et al., 2011), combining powerful visualization components (d3js, 2020) "to visually and interactively investigate transactional flow." The property that is pursued in an interactive visualization is the *multidimensional operability* and the *transparency*, the ability to show the internal complexity and to manage data autonomously.

Display Settings for Visual Reconstruction

In the comprehensive visual reconstruction of an organization, the convergence of data visualization and data analysis has become indispensable. The goal is to provide—in an interactive way—simultaneous calculation and visualization of the interconnected relationships among variables, distributions, and flow of processes in the different layers and phases of systems in organizations. "A useful starting point for designing advanced graphical user interfaces" is the Visual Information-Seeking Mantra of "seven tasks: overview, zoom, filter, details-on-demand, relate, history, and extracts (Schneiderman, 1996)" and to incorporate "the critical tasks that enable iterative visual analysis, including visualization creation, interactive querying, multiviewed coordination, history, and collaboration (Heer and Shneiderman, 2012). Visual analysis, modeling, and simulation of ecosystems and organizations are quite common, especially in the field of topological data analysis (Xu et al., 2017).

Complex adaptive systems modeling can be found in a wide range of areas from life sciences to networks and environments (CASModeling, 2020). Analysis and visualization of large networks can be performed with program packages, such as Pajek (Mrvar and Batagelj, 2016). The property that the configuration of an integrative visualization has to pursue is the *ubiquity* in order to accomplish a synthetic and holistic vision and analysis, which can be characterized as the capacity of understanding the complexity of a system by making it visible. The final step in the encoding of data visualization reaches the definition of the cross-layers of the functional system, which means to visually configure the vertical interconnection between the processes at their different layers.

Figure 3 shows a representation of the multilayered innovation ecosystem that involves science, technology, and business sub-ecosystems as an example of cross-layer analysis of collaborative network to investigate innovation capacities (Xu et al., 2017).

GRAPHIC DESIGN AND CONTEXT: MODAL APPROACHES AND PROPERTIES OF GOOD DATA VISUALIZATION

The fourth node of the communication framework is the context, which in data visualization is developed by graphic design. Data

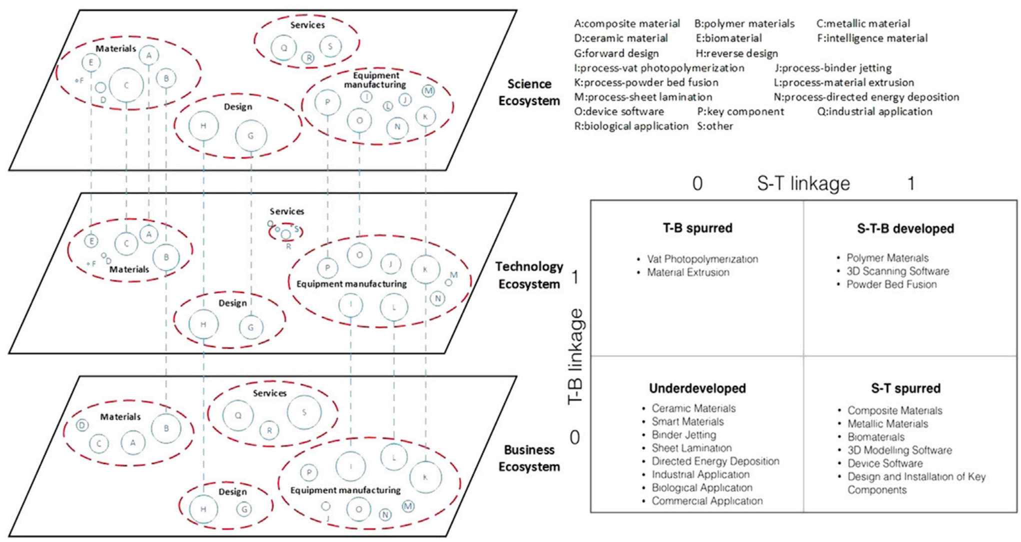

FIGURE 3 | Example of cross-layer analysis and visualization of a collaborative network in a science–technology–business ecosystem. Source: Xu et al. (2017).

visualization has to be adapted to the target and to the context where it is carried out through a suitable graphic design that captures the user's attention "to convey ideas effectively, both esthetic form and functionality" (Friedman, 2008). The effectiveness of the design of data visualization is evaluated by its impact on the user, and it is explained by the mechanisms of human perception of esthetic forms in particular contexts. "The most effective visualizations capitalize on the human facility for processing visual information, thereby improving comprehension, memory, and inference. Such visualizations help analysts quickly find patterns lurking within large datasets and help audiences quickly understand complex ideas" (Agrawala et al., 2011). The context is the criterion that classifies the approach modes to visualization and the esthetic forms of graphic design adopted.

Subjective Approach

Visualization must be meaningful. It has to pursue the properties of any communication act—clarity, concreteness, saving time, stimulating imagination and reflection, empowering the user, *etc.* For this purpose, from the perspective of computer-supported cooperative work (CSCW) studies, the idea of context and common ground being associated is important (Viégas and Wattenberg, 2006). In the subjective approach, the idea of context in its association with graphic design has to be defined considering the human–computer interaction (HCI). The principles of visual representation for screen design and the basic elements or resources used such as typography and text, maps and graphs, schematic drawings, pictures, node-and-link diagrams, icons and symbols, and visual metaphors should be observed. Engelhardt (2002) in his analysis of syntax and meaning in maps, charts, and diagrams establishes a classification of the correspondence systems between design uses and graphic resources (Blackwell, 2011). The principles of perception, visual processing, and the mechanisms and limitations of attention and memory, developed by the Gestalt School of Psychology—proximity, similarity, enclosure, closure, continuity and connection—define those mechanisms by which the human perception identifies patterns, forms, and organizations (Few, 2011), which explains that traditionally visualization reserves the spatial arrangement, the layout, for the most important dimensions of the data, to "code quantitative differences between objects and/or their relations" (Manovich, 2011). Complementing the coding that the brain automatically performs, the design can be used for recontextualization. The property that data visualization pursues through its graphic design in a subjective approach is *communicativity*, an essential condition or quality of being able to convey meanings from one entity or group to another through the use of mutually understood signs, symbols, and semiotic rules.

Objective Approach

Data visualization plays a critical role in multiple professional and academic fields, which means that it needs to adapt to particular specifications. The objective approach points to the context of professional specialization; for that reason, the graphic design must be basically functional in nature. Communication focuses on how to identify, instantiate, and evaluate domain-specific design principles for creating more effective visualizations (Agrawala et al., 2011). Graphic design is associated with graphic representation that can help the audience to understand better the relevant information. For instance, contour plots, heat maps, scatter-line combo, 3D graphs, or histograms can be especially useful in meteorology and environment, whereas line graphs, bar graphs, pie charts, mosaic or Mekko charts, population pyramids, and spider charts are usually more useful in marketing. Graphic design, to be effective, has to adapt to the functional needs in such a way that it has to modulate other principles of visualization. For instance, in Harry Beck's 1933 redesign of the London Underground station—because travelers only need to know the address and the remaining stops to reach their destination—aspects about the informational relevance of data were considered above others. From an objective approach perspective, the property that data visualization pursues through its graphic design is functional *adaptability*, a formal condition that refers to the ability to change in order to suit the needs of a new context or situation.

Informative Approach: Design as a Semiotic Mode

When there is an informative purpose, the communication effectiveness of the graphic design—beyond the subjective and objective approach that is constricted by the deontological reporting principle "to keep to the facts"—is mainly conditioned by the need to attract the attention of the user. "In an era of narrowly focused media that is often tailored toward audiences with a particular point of view, data visualization—and data journalism in general—offers the tantalizing opportunity for storytelling that is above all driven by facts, not fanaticism" (Cohen, 2020). The properties that graphic design of data visualization must meet in an informative approach can be assumed as properties of journalism. "Besides images, design is coming into play as a crucial semiotic mode for making meaning. In news features, special reports, or data visualizations, we can find a rich and complex interplay of different semiotic modes, for example, text, image, and layout, which constitute the meaning-making process" (Weber and Rall, 2016). In the field of data journalism, numerous examples of application of data visualization can be found, which are used to help to tell a story to readers (Cohen, 2020). In a fast-changing informational environment, graphic design in data visualization fundamentally has to be dynamic (Weber and Rall, 2016). The property that data visualization pursues through its graphic design in an informative approach is the *appealingness*, a modal condition of showing attractiveness that captures or awakes someone's interest.

Commercial Approach and Persuasive Communication

In the commercial approach, the graphic designer does not only try to capture the attention and interest of the user but also tries to convince him of the benefits of a product and a service. Visual communication can be fundamental as a complement of social influence. "Skilled visual designers manipulate the perception, cognition, and communication intent of visualizations by carefully applying principles of good design" that "can be used to either emphasize important information or de-emphasize irrelevant details" (Agrawala et al., 2011). Graphic design at this level is oriented to the presentation of a service, a concept, or a product, in which a clear persuasive intention is implied. Color choice, use of shapes, page layout, composition and focal points, rule of thirds, golden mean or divine proportion, eye path or visual hierarchy, balance, movement, white space, pattern, repetition, structure, type styling, grids, and alignment and contrast are effective design principles of persuasive communication that can make or break your marketing's effectiveness (Change Conversations, 2020). There are several types of graphic design that traditionally have been applied to marketing using "visual compositions to communicate ideas through typography, imagery, color, and form" such as visual brand identity, advertising, user interface, publication, packaging, environmental or way finding, art, and illustration (Cann, 2018). The property that data visualization pursues through its graphic design in a commercial approach is the *persuasivity*, an objective condition of being good at causing someone to do or believe something through reasoning or the use of temptation.

Educational-Investigative Approach

In contexts where learning or research processes take place, the design of data visualization is a factor of great importance. The synthesis and summary of data must be given in clear, attractive, and comprehensive graphic visualizations that show the logic of the internal connection of the elements or factors that participate in highly complex phenomena.

On the other hand, visualization requires user interaction, so the design has to adapt to the different phases of the learning or research process, or of discovery, be demonstrative, suggestive, progressive, *etc*. In the process of designing interactive visualizations for learning process, where performance, trial, and error are fundamental parts, in order to attempt to balance expressiveness, efficiency and accessibility visualizations can be greatly enhanced by interaction and animation (Bostock et al., 2011). Educational and scientific research approaches usually pursue synthetic graphical designs adapted to technical profiles. **Figure 4** shows "the UPC areas of knowledge and the relationship between them based on the scientific production of UPC researchers. Each node represents an area, and its size is determined by the number of activities in the portal FUTUR pertaining to it" (Future UPC, 2020).

Interactive visualizations are associated with techniques such as storytelling, which in turn are closely linked to graphic design. In her book "Design is Storytelling," Ellen Lupton reflects about the maxim "design is problem-solving" and how "designers use simple forms to convey compact messages" (Lupton, 2017). Plot, emotional connection, and simplicity (Less is more) have been described as three storytelling techniques for graphic design (Schauer, 2017). Investigative journalism is also one of the most important sources for producing interactive data visualization designs. The latest editions of the Online Journalism Awards (OJAs, 2019) or the Data Journalism Awards (2019) provide numerous examples of projects that allow interactive exploration. In an educational-investigative approach, the property that a data visualization pursues through its graphic design is the *dialogicity*, a demonstrative condition that takes the dialogue as "an efficient motivational strategy in encouraging participation in common efforts" (Zimmermann, 2011) for knowing the internal complexity, the detail of a case, a story, or an event.

Scientific Approach

The graphic design of data visualization in a scientific approach is a challenge that can be explained by different perspectives. From the point of view of collaborative experiences in applied research, it has been observed how "graphics are becoming increasingly important for scientists to effectively communicate their findings to broad audiences, but most researchers lack expertize in visual media" (Khoury et al., 2019). From the point

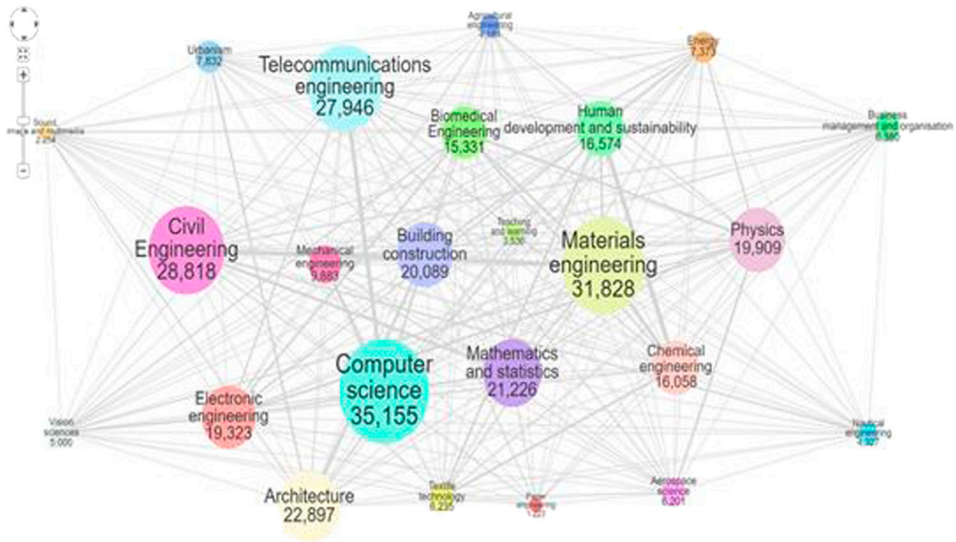

FIGURE 4 | Areas of knowledge and their relationship based on the scientific production of UPC researchers. Source: Future UPC (2020).

of view of scientists, "figures have a prominent role in scientific publications and often take up the majority of time when preparing a manuscript. Scientists and engineers would greatly benefit from having the appropriate design knowledge to draw effective figures" (Cheng and Rolandi, 2015). From a technological point of view, there are a large number of programs that provide solutions to support research and scientific communication, such as CartoDB (Carto, 2020) or Vizzuality (cfse, 2020). Data visualization experts point out that "good visualization is a winding process that requires statistics and design knowledge" (Yau, 2013). In its application to scientific dissemination, "well-constructed graphics can widen the impact of research articles" (Marcus, 2010; Khoury et al., 2019). Some initiatives have been proposed in order to maintain "key connection between the sciences and the visual arts," such as Design Help Desk, a project funded by the National Science Foundation that investigates the impact of visual design on scientific figures (Cheng and Rolandi, 2015). Finally, numerous companies in the field of visualization maintain a commitment to scientific dissemination and social responsibility associated with a vision that transcends the pragmatic use of visualization and data analysis (Periscopic, 2020). The property that data visualization pursues through its graphic design in a scientific approach is the *integrativity*, a condition of gathering in a visual unit the most detailed possible set of data and information of a complex reality with the possibility of interacting and experimenting with it.

MEDIA AND LEVELS OF COMMUNICATION EFFICIENCY

The fifth function of data visualization is to communicate relevant and objective information—understood as knowledge—in the most efficient way through the appropriate media. "The efficient communication of complex quantitative ideas" (Tufte, 1983) implicates the ability of being able to communicate successfully, minimizing the total resources of visualization taken in.

Content Editing: Correctness, Completeness, Timeliness, Accuracy, Review, and Control

The communication efficiency in editing the content of data visualization is measured in relation to its correctness, completeness, timeliness, accuracy, form, purpose, proof, and control. Statements in the "Disclaimer" section of websites related to the care and the accuracy with which the online content is created together with the rejection of any responsibility in case of incomplete or incorrect information are common (Human abilities, 2020). This major concern indicates the substantive contribution of the quality of the media diffusion of information on data visualization evaluation. Systems that verify the quality of the information have become extremely important. Not respecting this fundamental principle can lead to problems of social perception. For example, it is known that in the field of web-based social data analysis, "tools that rely on public discussion, to produce hypotheses or explanations of patterns and trends in data, rarely yield high-quality results in practice" (Willett et al., 2012). The control of data quality and its visualization are subject of study, for instance, in the framework of experiences and projects such as ESS Visio (2020) of the European Commission, where "sharing visualization tools between National Statistical Institute" has been successfully proposed (Laevaert and Le Goff, 2020). The basic characteristic that data visualization pursues through its media edition for diffusion is quality based on content *rigor*, an essential condition associated to *reliability* and verifiability that includes other characteristics—mentioned above—such as correctness or completeness.

Information Architecture and Navigability

Web navigability is a formal property that usually describes the ease with which a user moves through an "informational website"; therefore both the concept of navigability and the "guidelines for designing web navigation" for designing, managing, and augmenting effective link (Farkas and Farkas, 2000) applied to data visualization, and its diffusion in a communication channel, can be understood and assumed in similar terms. The *navigability* of data visualization can be conceptually examined along three dimensions: clarity of target, clarity of structure, and logic of structure (Wojdynski and Kalyanaraman, 2015). Effects such as reduction of search time, comprehension of content, and decrease of task time related to classical principles for the graphical design of interfaces such as "implications of memory, structure, and scent for information retrieval" (Larson and Czerwinski, 1998) must be equally considered in the dissemination of data visualization. Similarly, those contents related to the architecture of the information focused on the "structural design, organization, layout, and information" in the navigable space are applicable for data interactive visualization (SAP 2011; Uncharted Software, 2018; SAP 2020; Software, 2020). Associated with the concept of navigable space proposed by Lev Manovich (Manovich, 2001), the concept of mapping (Horn, 1998) has become one of the most prominent forms of visualization in the media associated with navigability as an exploratory activity—"Science mapping is a generic process of domain analysis and visualization" (Chen et al., 2015; Chen, 2017).

Technology and Visual Tools: Accessibility

Communication efficiency of data visualization is conditioned by accessibility which refers to the modal condition of being "usable by people with the widest possible range of abilities" (Henry et al., 2014). Applying the basic principles of *accessibility* that have been described as recommendations in the framework of the Web Content Accessibility Guidelines (WCAG 2.0), data visualization must have content that is "perceivable, operable, understandable, and robust" (W3C, 2019). Perceptibility as an efficiency factor has to be qualified on a quantitative level, in light of the well-known Weber–Fechner psychophysical law, according to which "the perceived change in stimuli is proportional to the initial stimuli" (Fechner, 1966). Perceptibility leads to the requirement that "information and user interface components must be presentable to users in ways they can perceive" (W3C, 2019), taking into account the ranking visualization of the correlation between stimuli and perception (Friendly, 2009; Demiralp et al., 2014; Kay and Heer, 2015; Prakash, 2016). Operations on a visualization interface allow the identification of salient patterns at various levels of granularity (Chen et al., 2015) in order to "promote comparison of terms both within and across latent topics" (Chuang et al., 2012) for assessing the data in context (Kandel et al., 2012). In the data-driven era, the understandability of the user interface is crucial to make timely decisions (Keim et al., 2013; W3C, 2019), in areas such as social media (Bello-Orgaz et al., 2016) or geoinformatics (Zuo et al., 2016). The ability to be "interpreted and managed reliably by a wide variety of user agents, including assistive technologies," which defines content robustness (W3C, 2019), has been observed in the field of "traffic data visualization" (Chen et al., 2015), social network (Hernández-Garcia et al., 2015), or designing animated transitions to convey aggregate operations (Kim et al., 2019).

Mediality: Multimedia, Hypermediality, and Multi- or Cross-Platform

Communication efficiency of data visualization is related to the concept of *mediality* which is the ability to appear in a communication medium, conditioned to understanding audiences, which "is integral to creating and distributing media messages" (Grady, 2020). "Multimediality is the interconnection of various functions which can provide media (text, images, graphics, animations, simulations, and so on)" (Brdička, 2003). The principle of multimediality is "used to be perceived as a way of facilitating the application of a wide range of transmission media" (Klement and Walat, 2015). In the field of education and especially in e-learning, multimediality "has to be understood as a means to stimulate multiple sides to pupil`s perception" (Klement and Walat, 2015). Hypermediality refers to digital content that, in addition to being in multimedia format, is interconnected in its configuration in order to facilitate navigation by user interaction. Hypertextuality refers to hypermediality restricted to the web publishing format.

Data visualization is closely related to multi-platform or cross-platform mediality forms "in the context of a wide range of distribution possibilities (e.g., online, mobile, and interactive games)" (Doyle, 2016). Multiples cross-platform data visualization solutions such as RGraph, Anychart, ZingChart, and DataGraph created by Visual Data Tools, such as Zoomcharts, are, among others, being developed by software companies. Criteria such as "avoiding repetition and increasing productivity" have been applied to assess cross-platform development approaches (Heitkötter et al., 2013). **Figure 5** shows an example of how the results of scientific research can be integrated in data journalism through innovative visualizations including multimedia contents being potentiality broadcasted in multiplatform media.

Media Interaction: Social Media, Cross-Mediality, and Data Storytelling

Communication efficiency of data visualization can be evaluated by its ability to assume different forms of interactive mediation between the user and the technology that gives access to the medium—"Interactive and dynamic graphics as part of multidimensional data analysis" (Cook and Swayne, 2007), "exploring high-dimensional data," and providing "highly dynamic and interactive graphics" (GGobi, 2020). "Visualization framework for real-time decision-making in a multi-input multi-output (MIMO) system" has been designed using statistical inferences in order to provide "accurate visual measures/ decision surfaces" (Ashok and Tesar, 2008). InSense, ManyEyes, and TweetPulse are some of the social big data applications that allow creating visualizations from collecting user experiences in collaborative environments through wearable data collection systems (Blum et al., 2006; Napalkova et al., 2018; WebLyzard, 2020). Data visualization is a key technical challenge when designing Cross Media games, employing "a wide variety of

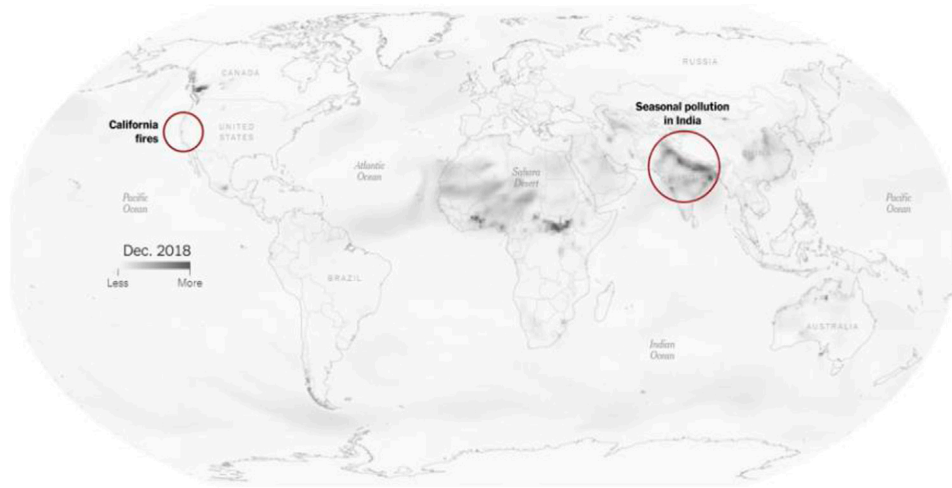

FIGURE 5 | Screen capture of the Data Journalism Award (2019) Best visualization (large newsrooms). Winner: See How the World's Most Polluted Air Compares with Your City's. Organization: The New York Times. https://www.nytimes.com/interactive/2019/12/02/climate/air-pollution-compare-ar-ul.html.

gaming interfaces based on stationary and mobile devices to facilitate different game experiences within a single game" (Lindt et al., 2005, Lindt et al., 2007) such as the alternate reality games (Walz, 2010). The evaluation of the efficiency of data visualization is also related to its capacity for transmediality, where consumers play an active role in different platforms and media (Chen et al., 2015). Investigative journalism also incorporates the concept of data storytelling or data narrative where ideas must be supported by data while maintaining and demonstrating rigor in their processing. Elements that participate in the narrative according to info graphic taxonomies have been categorized (Ruys, 2020). In the last decade, publications on the convergence of data visualization and data storytelling are experiencing rapid growth (Segel and Heer, 2010; Hullman et al., 2013; Kosara and Mackinlay, 2013; Knaffic, 2015; Lee et al., 2015). The *multimedia interactivity* or *participativity* is the ability to promote interactive access to users in order to spread a message or a story, a demonstrative condition that can be used to measure the communication efficiency of data visualization once it is projected in the media.

Meta-Mediality: Augmented Reality, Hyperreality, and Mixed Reality

Once "life coaching has been presented as a collaborative social action of storing and sharing users life events in an open environment" (Bello-Orgaz et al., 2016), the next step is to enable visualization to recreate a reality from the media but especially beyond the media. "What happens to the idea of a 'medium' after previously media-specific tools have been simulated and extended in software? Is it still meaningful to talk about different mediums at all? Or do we now find ourselves in a new brave world of one single monomedium, or a metamedium" (Manovich, 2013, *borrowing Kay's term*). Metamediality, applied to data visualization, can be understood as a transcendental condition in as much as its aim is to overcome the figure of the medium as intermediary, seeking to transcend the reality that it explains, creating a new one (Kay and Goldberg, 1977). Metamediality can be understood as a mix between metafiction and intermediality, ranging from augmented reality (AR) as an interactive experience, and hyperreality, where consciousness is unable to distinguish reality from a simulation, to mixed reality (MR) as the merging of real and virtual worlds to produce new environments and visualizations, where physical and digital objects coexist and interact in real time. The possibility of recreating and living the data visualization by the user constitutes a transcendental capacity of *experiment-ability* that defines data visualization when it transcends the medium where it is projected.

USER AND USABILITY REQUIREMENTS

In the visualization process and as a culmination of it, the requirements arising from the interaction and user experience must be considered, which are defined as components of usability. Usability is a "quality attribute that assesses how easy user interfaces are to use" and "also refers to methods for improving ease-of-use during the design process (Nielsen Norman Group, 2012). Some of these requirements are included in the ISO 9241-11, "The objective of designing and evaluating systems, products, and services for usability is to enable users to achieve goals effectively, efficiently and with satisfaction, taking account of the context of use" (ISO, 2020). These goals have been proposed as components of usability as parameters to be tested: learnability, efficiency, memorability, errors, and satisfaction" (Nielsen Norman Group, 2012).

Perception: Learnability and Flexibility

The essential requirement that results from the user experience of data visualization is learnability, "how easy is it for users to

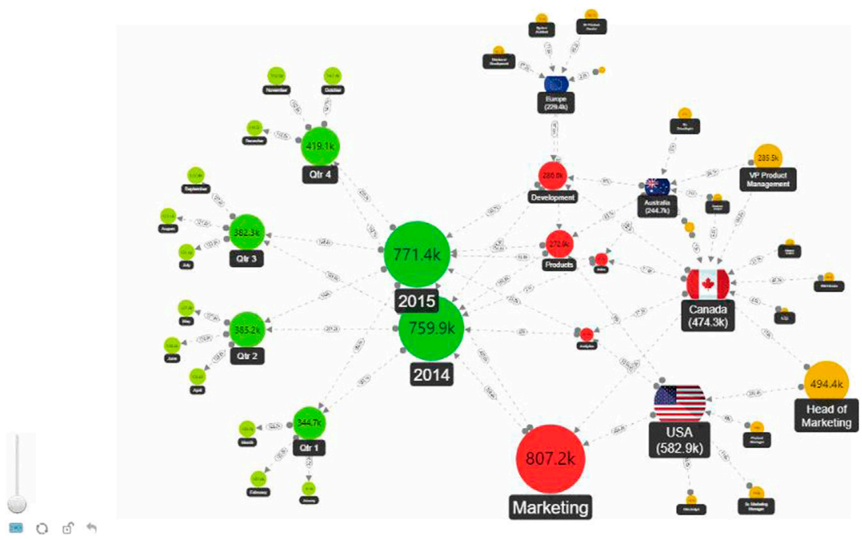

FIGURE 6 | Screen capture of interactive ZoomCharts Network Chart Custom Visual for Microsoft Power BI which is supported on all platforms and can be implemented and customized by the user (ZoomChart, 2020).

accomplish basic tasks the first time they encounter the design? (Nielsen Norman Group, 2012) Learnability in data visualization can be defined as the attribution or basic quality necessary to enable a user to learn from it and learn to interact with it. To this end, "manual and automated chart specification to help analysts engage in both open-ended exploration and targeted question answering" have been developed (Wongsuphasawat et al., 2017) in order to facilitate the user experience: how people should use the information, what they should use it to accomplish" (Few, 2011). The *learnability* requirement leads to *flexibility*, the ease of operating interactively with all the possibilities with which the user and the system can exchange information: possibility of dialogue, multiplicity of ways to carry out the task, and similarity with previous tasks (Fernández Casado, 2018), ensuring user familiarity with visual analysis tools (Kim et al., 2019).

The network chart shown in **Figure 6** illustrates how innovative software and applications are leading a new open approach for data visualization that allows the user to customize the parameters of their preferences according to their own criteria.

Accomplishment of Tasks and Efficiency

The second level of the user experience in the use of data visualization occurs when the user is active and has an autonomous experience. The consequent question formulated in terms of *efficiency* as a component of web usability is the following: "Once users have learned the design, how quickly can they perform tasks?" (Nielsen Norman Group, 2012). Here, efficiency is evaluated as a formal requirement of usability that is defined in terms of resources, such as time, human effort, costs, and materials (ISO, 2020) deployed for the accomplishment of tasks. Efficiency related to the accomplishment of tasks is a requirement that can be evaluated by observing the quality of the autonomous experience that the user has when using and interacting with the visualization. Efficiency in usability can be measured based on performance data–applying methods, similar to Ads Quality Score that is obtained by analyzing the relevance of the content, the loading speed, the quality, and relationship of the images, texts, links, *etc.* (Google, 2020). Obtaining major detail about the relation between user performance and experience is possible. For instance, in life-logging services, different factors of user experience are recorded. Sensors can capture continuous physiological data, such as "mood, arousal, and blood oxygen levels together with user activities" (Bello-Orgaz et al., 2016). The usability analysis can be supported by techniques of "exploratory analysis of dynamic behavior of individuals visiting a particular website" (Cadez et al., 2003).

Performance and Effectiveness of Data Visualization in Knowledge Transfer

When the user operates with data visualization as an instrument for the representation, analysis, or visual communication of data the user's actions are oriented toward how visualization facilitates the process of transferring information and knowledge so that efficiency becomes effectiveness, as a modal requirement of performance for the user that must be evaluated. The *effectiveness* as a measurable element can be defined as the "accuracy and completeness with which users achieve specified goals" (ISO, 2020). A classic example of communication effectiveness can be observed in the famous *Anscombe quartet* that Edward Tufte used to illustrate the importance of visualization as an instrument of analysis and therefore for the transfer of knowledge (Tufte, 1983). The effectiveness and *performance* as a component of the usability of a visualization comes from the use that visualization makes of the resources of the human visual system as a processor to detect patterns, trends, or anomalies, which explains the use of facilitating plugins based on perceptive factors. For example, in the field of "Designing

Animated Transitions to Convey Aggregate Operations," recent studies "indicate that staged animated transitions can improve the subject's ability to correctly identify aggregation operations, although sometimes with longer response times than with static transitions (Kim et al., 2019).

Proficiency and Memorability

A higher level of complexity in the requirements for good visualization based on the user experience is reached when the user is empowered by the acquired knowledge and expert mastery of the visualization tool. Here, the requirement that visualization must achieve is to enable the user to improve his experience by incorporating his own contributions or preferences, expanding the framework of action, and applying this experience to other cases. Here, it is necessary to consider the competencies of the user in relation to the configuration of the human brain, which in turn corresponds to the different dimensions of the human as a self-conscious being. When the user operates proactively, the spatial working memory (SWM) "plays an essential role in driving high-level cognitive abilities," and it "is associated with global brain communication" (Liu et al., 2017). *Memorability* as a component of usability can be considered in terms of determining how fast and easy it is for the user to "reestablish proficiency" (Nielsen Norman Group, 2012) and has been developed in relation to visualization recognition and recall "what components of data visualization attract people's attention, and what information is encoded into memory" (Borkin. et al., 2015). As in the case of effectiveness, reestablishing *proficiency* can be improved in an assisted manner. For example, "animation can help viewers track changes and stay oriented across transitions between related statistical graphics with research to date primarily focused on transitions in response to filtering, time steps, changing variables, or adjusting visual encodings (Kim et al., 2019).

Feedback, Interaction, and Error Prevention: Supportiveness and Robustness

The evaluation of the usability of data visualization tools can be carried out by studying the errors made by the user with the objective of introducing improvements for future prevention and for enhancing their *robustness*. The general question is: "How many errors do users make, how severe are these errors, and how easily can they recover from the errors?" (Nielsen Norman Group, 2012). The *supportiveness* is a requirement that seeks to empower the user for his success through training services, help, support consultation generated by self-learning automated systems that identify and correct errors, and irregularities. Applied to data visualization tools, such as Lyra, this ability has been studied in association with their interactive capacity (Satyanarayan and Heer, 2010). Interactive visualizations have been incorporated into the design of applications in the context of machine comprehension based on error analysis, for example, in NLP (natural language processing) such as Errudite (Wu et al., 2019). In order to understand the user's participation in the content of the visualization, making them part of a social process and a learning community, there are tools developed to help users in order to obtain better visualizations. Mixed-initiative systems such as Voyager have been designed in order to support "faceted browsing of recommended charts chosen according to statistical and perceptual measures" (Wongsuphasawat et al., 2017).

Global Experience and Satisfaction

The user's experience with data visualization is summarized in the satisfaction obtained; so the study of the requirements of data visualization culminates with the explanation of the contributing factors. Determining how attractive the design is to use (Nielsen Norman Group, 2012) "can make a difference on whether or not users come back to it" (Chrisdasie, 2020). Satisfaction refers to the "extent to which the user's physical, cognitive, and emotional responses that result from the use of a system, product, or service meet the user's needs and expectations" (ISO, 2020). In the evaluation of visual communication, it has been proposed to obtain early feedback on the level of user satisfaction through questionnaires or qualitative interviews, as well as through analytics of the use of visualization and other more sophisticated techniques such as movement analysis eyes when users use visualization (Agrawala et al., 2011). There are studies on usability and the user satisfaction of hardware–software interfacing visualization that have demonstrated the need to develop educational research on the use of display technologies, such as in the field of learning programming (Ali and Derus, 2014). User *satisfaction* is a requirement that has prompted a large number of studies in the scientific literature, some of which have even proposed the development of an "ontology visualization tool, to provide a user-centered interactive solution" for extracting and visualizing Linked Data (Ghorbel et al., 2017). Experimental evidence indicates that research on systems for evaluating the degree of accomplishment of data visualizations is still incipient. In similar terms that can be stated about the certainty of scientific theories, user's satisfaction cannot be certified, but dissatisfaction can eventually be demonstrated.

RESULTS

The results of the study conducted in this article can be classified into two groups: *theoretical*—which include (a) dimensional factors, (b) characterization of achievements—and *practical*, which include (c) types of data visualization, (d) functions, (e) principles of assessment, and (f) professional competences of data visualization.

a) DIMENSIONAL TAXONOMY OF DATA VISUALIZATION. **Table 2** shows the dimensional taxonomy with indication of the factors of completeness and complexity for each stage of procedure and progress of data visualization.
b) CHARACTERIZATION OF ACHIEVEMENTS. The nature of the conditions or properties in the procedure of data visualization follows a common pattern of a sequential order. **Table 3** shows the following: in the basic layer, substantial or essential conditions that must be achieved by data visualization; in the extended, formal conditions; in the synthetic, modal conditions; in the

TABLE 2 | Dimensional taxonomy of data visualization: factors of completeness and factors of complexity.

		Factors of completeness					
	Elements	**Content**	**Graphic Representation**	**Encoding Set-up**	**Graphic design**	**Media**	**User**
	Dimensions	**Degrees** of abstraction	**Functionalities** of the tools for the graphical representation	**Specifications** of the set-up of the visualization	Approach **modes** and **properties** of visualization	**Levels** of communication efficiency	**Requirements** from the user experience side
Factors of complexity	**Layers**						
	Basic	Parameters and scales	Descriptive graphs	Formalization and basic setup	Subjective	Edition	Perception
	Extended	Indicators and relations	Multivariate or relational graphs	Transformation	Objective	Information architecture	Accomplish-ment of tasks
	Dynamic	Variable sets distribution	Dynamic graphs	Integration	Informative	Technology and visual tool	KT performance
	Synthetic	Processes and phases	Process info graphs and motion graphics	Modelization of processes	Commercial	Mediality: multimediality	Proficiency
	Interactive	Hyper-processes layers and systems	Interactive graphics	Interactive visual analysis	Educational	Media interaction	Feedback and errors
	Integrative	Organizations and ecosystems	Scientific graph	Ecosystem modeling	Scientific	*Meta*-mediality	Global experience

dynamic, objective conditions; in the interactive, demonstrative conditions; and finally in the integrative layer, transcendental conditions. From a practical point of view, the design of a dimensional taxonomy of data visualization may cast fresh light on the types, functions, principles, and required competences for data visualization.

c) TYPES OF VISUALIZATION. Once an object-centered model of data visualization has been defined—as previous exploratory and experimental studies have shown (Cavaller et al., 2020)—six types of data visualization can be obtained (see **Table 4**).

d) FUNCTIONS OF DATA VISUALIZATION. According to the defined taxonomy, factors, and achievements, the functions of data visualization are the following (see **Table 5**):

1. The first function of data visualization is to show the relationship among the parameters that describe a phenomenon, a process, a system, or any observable subset of the real world.
2. The second is to represent data in a visual way by a graphical representation.
3. The third function is to communicate, that is, to convey meaning—transforming data into information—to be understood by someone.
4. The fourth function is the dissemination of a meaning content by a graphic design appropriate to the context where it will be communicated.
5. The fifth function is to communicate relevant and objective information in the most efficient way through the appropriate media.
6. The sixth function of data visualization is to observe the restraints, capabilities, and conditions from the users in order to enhance the communication performance.

e) PRINCIPLES FOR THE ASSESSMENT OF DATA VISUALIZATION. Data visualization can be assessed according to six different principles of interests.

1. The principle of analytical interest states that data visualization is right in so far as it keeps scientific rigor, order, and method in the quantitative procedures

TABLE 3 | Dimensional taxonomy of data visualization: properties or conditions of data visualization.

	Visualization element	**Content**	**Graphical representation**	**Encoding setup**	**Graphic design**	**Media**	**User**
Achievement	**Essential**	Congruence	Evidentiality	Accuracy	Communicativity	Reliability	Learnability and flexibility
	Formal	Exhaustivity	Proportionality and integrity	Multi-directionality	Adaptability	Navigability	Efficiency
	Modal	Consistency	Schematicity	Versatility	Appealingness	Accesibility	Effectiveness and performance
	Objective	Cohesive unity	Flow logicality	Self-explanatority	Persuasivity	Mediality	Proficiency and memorability
	Demonstrative	Falsibiability	Complex evidentiality	Operatibility and transparency	Dialogicity	Participativity	Supportiveness and robustness
	Transcendental	Modeling	Intricacy	Ubiquity	Integrativity	Experiment-ability	Satisfaction

TABLE 4 | Variables, types of visualization, and graphical representation by goals from the perspective of an object-centered data visualization model (Cavaller et al., 2020).

	Content-variable	Types of data visualization and graphical representation	Object-goal
1	Measurements	Descriptive	Parameters and basic relationships
2	Indicators	Relational	Multivariate relationships
3	Distributions	Multi-relational dynamics	Factors or multi-relationships
4	Flow: vector	Process	Internal logics
5	Network: connector	Hyper-process: system	Architecture
6	Program: code	Ecosystem	Organization

TABLE 5 | Taxonomy of data visualization: functions, principles, and competences in data visualization.

Visualization element	Content	Graphical representation	Setup	Graphic design	Media	User experience
Component	**Message**	**Form**	**Encoder**	**Context**	**Channel**	**Decoder**
Functions	Show parametrical relations	Represent data in a visual way	Convey meaning	Dissemination of information	Efficient communication	Enhance communication performance
Principles	Analytical	Functional	Managerial	Efficacy	Efficiency	Appraisal
Competences	Data analysis	Data graphic representation	Programming	Data graphic design	Media publishing	Human–machine interaction

2. The principle of functional or pragmatic interest states that data visualization is right in so far as the graphical representation has a practical utility and added value over other communicative forms facilitating their comprehension.
3. The principle of managerial interest states that data visualization is right in so far as it is able to package data-message and graphic representation in a singular configuration that promotes the understanding of a meaningful communication.
4. The principle of interest for efficacy states that data visualization is right in so far as, taking into account the professional, social and cultural context and target; it produces the intended communicative result by a suitable design.
5. The principle of interest for efficiency states that data visualization is right in so far as it achieves the communication goals by the optimal means of communication with maximum benefits and minimal use of resources
6. The principle of appraisal interest states that data visualization is right in so far as it receives a positive assessment from the user in terms of usability and of other factors related to H–M interaction.

f) According to the functions and principles mentioned above, data visualization can be defined as a multidisciplinary field where professionals need a wide range of knowledge specializations and professional competences such as data analysis, data graphic representation, programming, graphic design, media publishing, and human–machine interaction.

FINDINGS AND DISCUSSION

The fundamental conceptual findings of the study include the following:

1) The process of data visualization can be viewed from the perspective of communication sciences, which includes six major components: message, form, encoder, context, channel, and decoder.
2) These components correspond to six elements in the context of data visualization: content, graphic representation, encoding setup, graphic design and approach, media, and user.
3) The process of data visualization must integrate complexity as a parameter in its implementation, and it must be ordered

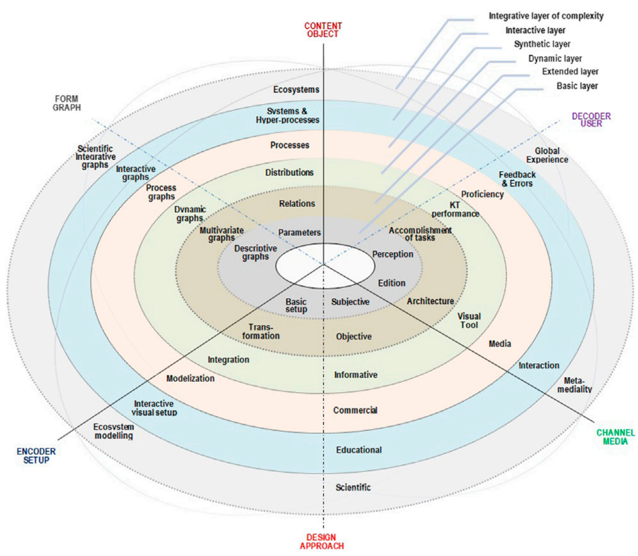

FIGURE 7 | Illustrative representation of the dimensional taxonomy for object-oriented data visualization from the perspective of communication sciences: elements-axes as factors of completeness and layers spheres as factors of complexity. Source: Own elaboration.

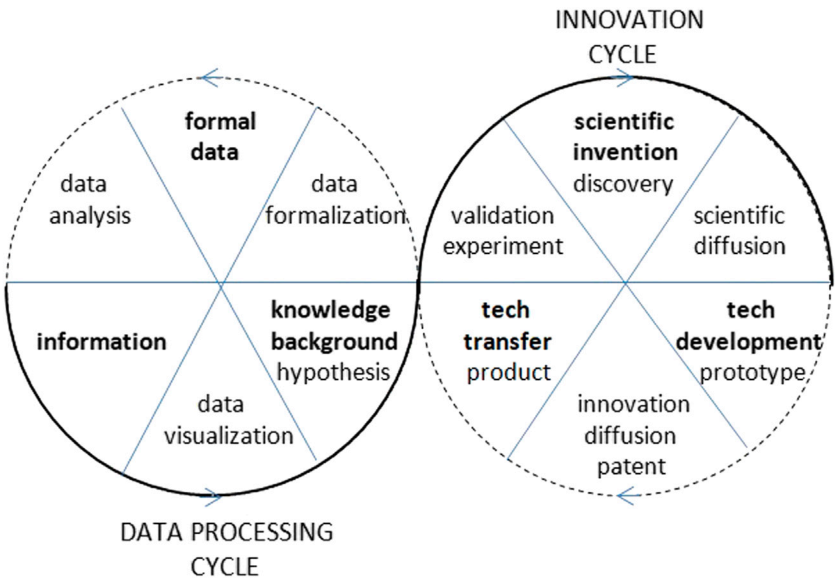

FIGURE 8 | Representation of the sequentiality of data processing and innovation hyper-cycle.

according to six layers of complexity: basic, extended, dynamic, synthetic, interactive, and integrative. These layers, obtained by analytical criteria, indicate the degree of the internal complexity of the organized entity or a phenomenon that is represented, and they are defined in order to facilitate the systematic application of object-oriented data visualization.

The process of data visualization must be addressed following the unfolding of the possibilities that arise from the combination of these factors, reaching the observed achievements at each crossroads between *communication component* x *layer of organizational complexity* (see **Figure 7**).

Previous theoretic and practical studies have led to the assumption that data visualization is mainly instrumental. Conversely, the results of this study reveal that the potentialities of the analytical functions of data visualization are strictly related to its ability to show the scale and the increasing intricacy of the networked organization of a complex system, in which relationships and processes are interconnected.

In other terms, the efficacy of data visualization not only depends on the completeness of its *extended deployment* taking into account communication factors but also on its *in-depth unfolding* following the level of organizational complexity in which the analysis has been performed. This holistic approach enables data visualization to be understood as the visual representation of knowledge, after data formalization and data analysis. As the key time that culminates and completes data processing, data visualization summarizes the underlying background knowledge that potentially initiates a new inquiry in the innovation cycle.

For an open discussion, it must be pointed out that the completion of data visualization, according to the proposed taxonomy, culminates data processing cycle, making visible the knowledge background. On this basis, scientific research, technological development, and transfer deploy the cycle of innovation (Cavaller, 2008), which, in turn, pushes data processing cycle for the extension of scientific knowledge (see **Figure 8**). So, in a major hyper-cycle, data processing and innovation cycles can be seen as an augmented projection of human cognitive process, where this taxonomy of data visualization can play an extended key role, an issue that constitutes the object for future research actions.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTION

The author confirms being the sole contributor of this work and has approved it for publication.

ACKNOWLEDGMENTS

I thank Ricardo García (UOC) and Anahí Martin (El Arte de Medir) for providing the visualization of **Figure 1**.

REFERENCES

Agrawala, M., Li, W., and Berthouzoz, F. (2011). Design principles for visual communication. *Commun. ACM* 54 (4), 60–69. doi:10.1145/1924421.1924439

Ali, A. Z. M., and Derus, S. R. (2014). Hardware-software visualization to learn programming. The 9th International Conference on Computer Science & Education (ICCSE 2014). Vancouver, Canada. doi:10.1109/ICCSE.2014.6926508

Allen, T. F. H., and Starr, T. B. (2017). *Hierarchy: perspectives for ecological complexity*. London. University of Chicago Press.

Ambicon (2016). Final Report Summary - AMBICON (Automated multi-level modeling of biological systems considering physico-chemical constraints). Available at: https://cordis.europa.eu/project/id/332020/reporting.

Aparicio, M., and Costa, C. J. (2014). Data visualization, *Commun. Des. Q.* 3. 1. doi:10.1145/2721882.2721883

Ashok, P., and Tesar, D. (2008). A visualization framework for real time decision making in a multi-input multi-output system. *IEEE Syst. J.* 2. doi:10.1109/jsyst.2008.916060

Barnlund, D. C. (2008). "A transactional model of communication," in *Communication theory*. Editor C. D. Mortensen. 2nd ed. (New Brunswick, New Jersey: Transaction.), pp 47–57.

Bello-Orgaz, G., Jung, J. J., and Camacho, D. (2016). Social big data: recent achievements and new challenges. *Inf. Fusion* 28, 45–59. doi:10.1016/j.inffus.2015.08.005

Benoît, G. (2019). *Introduction to information visualization. Transforming data into meaningful information*. Newyork, London, Rowman & Littlefield. 61, 300–302. doi:10.3138/jelis.2019-0026

Berlo, D. K. (1960). *The process of communication*. New York, New York: Holt Rinehart & Winston.

Bertin, J. (1967). (re-ed. **2011**). *Semiology of graphics*. London, United Kingdom: Esri Press.

Blackwell, A. (2011). *Visual representation». The encyclopedia of human-computer interaction* (2a. ed., cap. 5).

Blum, M., Pentland, A., and Troster, G. (2006). InSense: interest-based life logging. *IEEE MultiMedia* 13 (4). doi:10.1109/mmul.2006.87

Borkin, M. A., Bylinskii, Z., Kim, N. W., Bainbridge, C., Yeh, C. S., Borkin, D., et al. (2015). Beyond memorability: visualization recognition and recall. *IEEE Trans. Vis. Comput. Graph* 22 (1), 519–528. doi:10.1109/TVCG.2015.2467732

Börner, K., and Polley, D. E. (2014). *Visual insights: a practical guide to making sense of data*. London. MIT Press.

Börner, K. (2015). *Atlas of knowledge: anyone can map*. London. MIT Press.

Bostock, M., Ogievetsky, V., and Heer, J. (2011). D^3: data-driven documents. *IEEE Trans. Vis. Comput. Graph* 17 (12), 2301–2309. doi:10.1109/TVCG.2011.185

Boyer, S. A. (2010). *SCADA: supervisory control and data acquisition*. Fourth Edition. Durham, United States: ISA.

Brdička, B. (2003). *Role Internetu ve vzdělávání: studijní materiál pro učitele snažící se uplatnit moderní technologie ve výuce*. Kladno: AISIS.

Cadez, I., Heckerman, D., Meek, C., SmythWhite, P. S., and White, S. (2003). Model-based clustering and visualization of navigation patterns on a web site. *Data Mining Knowledge Discov.* 7 (4), 399–424. doi:10.1023/a:1024992613384

Cairo, A. (2013). *The Functional Art: an introduction to information graphics and visualization*. New Riders.

Cann, M. J. (2018). The 8 types of graphic design. 99designs. Available at: https://99designs.es/blog/tips/types-of-graphic-design/.

Card, S. K., Mackinlay, J. D., and Schneiderman, B. (1999). *Readings in information visualization: using vision to think*. San Francisco: Morgan Kaufmann Publishers.

Carto (2020). Unlock the power of spatial analysis. Available at: https://carto.com/downloads/carto-press-kit.pdf (Accessed April 2, 2021).

CASModeling (2020). Complex adaptative systems modeling. Available at: https://casmodeling.springeropen.com/.

Cavaller, V., García, R., and Martin, A. (2020). "Knowledge transfer process visualization: dealing with complexity" ICICKM Procceedings. 17th International Conference on Intellectual Capital A Virtual Conference. (2020). Kidmore End, United Kingdom: Knowledge Management & Organisational Learning.

Cavaller, V. (2008). Matrix indicator system. Collnet Fourth International Conference on Webometrics, Informetrics and Scientometrics (WIS) & Ninth COLLNET Meeting, July -1 August, 2008, Berlin, Germany. 28 Available at: http://www.collnet.de/Berlin-2008/CavallerWIS2008mis.pdf.

Cavaller, V. (2007). Matrix indicator system for strategic information analysis: application in a practical case Dissertation. Ph.D. *In 'information and documentation in digital age' program*. Barcelona, Spain: the Department of Library and Information Science at the University of Barcelona. Available at: https://www.tdx.cat/handle/10803/763#page=1.

Cavaller, V. (2020). *"Multidimensional assessment of knowledge transfer in higher education: the Sextuple Helix Model for science and society progress and the KT Quality Index" ICICKM Procceedings* 17th International Conference on Intellectual CapitalA Virtual Conference. 15-16 October 2020. Kidmore End, United Kingdom: Knowledge Management & Organisational Learning.

Cfse (2020). Vizzuality**.** Cambridge forum for sustainability and the environment. Available at: https://www.cfse.cam.ac.uk/directory/craig_mills.

Chen, C. (2017). Science mapping: a systematic review of the literature. *J. Data Inf. Sci.* 2 (2), 1–40. doi:10.1515/jdis-2017-0006

Chen, W., Guo, F., and Wang, F.-Y. (2015). A survey of traffic data visualization. *IEEE Trans. Intell. Transport. Syst.* 16, 2970. doi:10.1109/TITS.2015.2436897

Cheng, K., and Rolandi, M. (2015). Graphic design for scientists. *Nat. Nanotechnol* 10, 1084. doi:10.1038/nnano.2015.290

Chrisdasie (2020). 5 components of usability. Available at: https://chrisdasie.com/5-components-of-usability/#.

Chuang, J., Manning, C. D., and Heer, J. (2012). Termite: visualization techniques for assessing textual topic models. Proceedings of the international working conference on advanced visual. May 2020. doi:10.1145/2254556.2254572

Cohen, S. (2020). UsingvisualizationtotellstoriesDataJournalism.com. Available at: https://datajournalism.com/read/handbook/one/delivering-data/using-visualizations-to-tell-stories. doi:10.1287/ad139370-f916-4a8b-bb9d-c880275f72c7

Change Conversations (2020). 10 effective design principles you need to know. Available at: https://www.marketing-partners.com/conversations2/10-effective-design-principles-you-need-to-know.

Cook, D., and Swayne, D. F. (2007). *Interactive and dynamic graphics for data analysis: with R and GGobi*. USA. Springer.

Curcin, V., Ghanem, M., and Guo, Y. (2010). The design and implementation of a workflow analysis tool. *Philos. Trans. A. Math. Phys. Eng. Sci.* 368, 4193–4208. doi:10.1098/rsta.2010.0157

D3js (2020). Available at: https://d3js.org/.

Datajournalismawards (2019). Available at: https://datajournalism.com/awards.

DeLanda, M. (2019). *A new philosophy of society: assemblage theory and social complexity*. London.

Demiralp, Ç., Bernstein, M. S., and Heer, J. (2014). Learning perceptual kernels for visualization design. *IEEE Trans. Vis. Comput. Graph* 20 (12), 1933. doi:10.1109/TVCG.2014.2346978

Doyle, G. (2016). Multi-Platform Media: has digitization really given us more for less? EconomistTalkArt.org. Available at: https://www.digitalmeetsculture.net/article/multi-platform-media-has-digitization-really-given-us-more-for-less/.

Engelhardt, Y. (2002). The Language of Graphics. A framework for the analysis of syntax and meaning in maps, charts and diagrams. (PhD Thesis). Amsterdam (Netherlands): University of Amsterdam.

Farkas, D., and Farkas, J. B. (2000). Guidelines for designing web navigation. *Tech. Commun. (Washington)* 47 (3), 341–358.

Fayyad, U., and Uthurusamy, R. (2002). Evolving data into mining solutions for insights. *Commun. ACM* 45 (8), 28–31.

Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Mag.* 17 (3), 37.

Fechner, G. Th. (1966). *Elements of psychophysics [elemente der Psychophysik*. (New York, Holt, Rinehart and Winston). [First published. 1860. Editors D. H. Howes and E. G. Boring, Vol. 1.

Fernández Casado, P. E. (2018). *Usabilidad web*. Madrid, Spain, Teoría y Uso. Rama.

Few, S. (2011). Data visualization for human perception. The encyclopedia of human computer interaction. Interaction design foundation. Available at: https://www.interaction-design.org/.

Friedman, V. (2008). Data visualization and infographics. *Smashing Mag. January* 14. doi:10.5912/jcb231

Friendly, M. (2006). *A brief history of data visualization». Handbook of Data Visualization,*Springer, Berlin, Heidelberg., Heidelberg, Germany, 15–56. doi:10.1007/978-3-540-33037-0_2

Friendly, M. (2009). *Milestones in the history of thematic cartography, statistical graphics, and data visualization*. Corpus ID: Computer Science., 64305987.

Future UPC (2020). Website for the scientific production of UPC researchers. Available at: https://futur.upc.edu/.

GGobi (2020). GGobi. Out of sight, out of mind. Available at: http://ggobi.org/.

Ghorbel, F., Hamdi, F., Ellouze, N., Métais, E., and Gargouri, F. (2017). Visualizing large-scale linked data with memo graph, *Acad. J. Acad. J. |Proce. Computer Sc.* Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 21st International Conference, KES-2017, 6-8 September 2017, Marseille, France, 112, 854–863.

Gibb, S., Hendry, R. F., and Lancaster, T. (2019). *The Routledge Handbook of Emergence*. New york, Routledge.

GIS (2020). 25 map types: brilliant ideas to build unbeatable maps. Available at: https://gisgeography.com/map-types/.

Goddar, T. D., Huany, C. C., Meng, E. C., Pettersen, E. F., Couch, G. S., Morris, J,H., et al. (2018). UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* 27 (1), 14–25. doi:10.1002/pro.3235

Google (2020). About quality score. Available at: https://support.google.com/google-ads/answer/7050591.

Grady, D. (2020). *The golden age of data: media analytics in study & practice*. New York. Routledge.

Greenberg, I., Xu, D., and Kumar, D. (2013). *Processing: creative coding and generative art in processing 2*Greenberg, Ira., Xu, Dianna., and Kumar, Deepak. Friends of ED, New York.

Harary, F. (1969). *Graph theory*. Addison-Wesley series in mathematics.

Harrison, L., Reinecke, K., and Chang, R. (2015). "Infographic aesthetics: designing for the first impression". CHI 2015: Proceedings of the 33rd Annual CHI Conference on Human Factors in Computing Systems. Seoul, South Korea, April 18-23.

Hausmann, R., Hidalgo, C. A., Bustos, S., and Coscia, M. (2014). *The atlas of economic complexity: mapping paths to prosperity*. MIT Press.

Heer, J., and Agrawala, M. (2006). Software design patterns for information visualization. *IEEE Trans. Vis. Comput. Graph* 12 (5), 853. doi:10.1109/TVCG.2006.178

Heer, J., and Shneiderman, B. (2012). "Interactive Dynamics for Visual Analysis: a taxonomy of tools that support the fluent and flexible use of visualizations". *ACM Queue* 10. 2, doi:10.1145/2133416.2146416

Heitkötter, H., Hanschke, S., and Majchrzak, T. A. (2013). in *Evaluating cross-platform development approaches for mobile applications. Webist 2012, lnbip 140*. Editors J. Cordeiro and K.-H. Krempels (Springer-Verlag Berlin Heidelberg.), 120–138.

Henry, S. L., About-Zahra, S., and Brewer, J. (2014). The role of accessibility in a universal web. Proceedings of the 11th Web for All Conference. Cambridge, MA 02139-4307 USA. April 2014. doi:10.1145/2596695.2596719

Hernández-García, Á., González-González, I., Jiménez-Zarco, A. I., and Chaparro-Peláez, J. (2015). Applying social learning analytics to message boards in online distance learning: a case study. *Comput. Hum. Behav.* 47, 68–80. doi:10.1016/j.chb.2014.10.038

Hofffman, P., Grinstein, G., Marx, K., and Grosse Stanley, I. E. (1997). DNA visual and analytic data mining, *Proc. Visualization*. 97, 437–441.

Hofffman, P., Grinstein, G., and Pinkney, D. (1999). Dimensional anchors: a graphic primitive for multidimensional multivariate information visualizations. *Proceedings of the 1999 workshop on new paradigms in information*. Kansas City, Missouri, USA. 1–19. doi:10.1145/331770.331775

Horn, R. E. (1998). *Visual language: global communication for the 21st century*. WA: Bainbridge IslandMacroVU, Inc.

Hu, B., JinGuo, J. A. Y., Zhang, H., Luo, J., and Gao, G. (2014). Gsds 2.0: an upgraded gene feature visualization server. *Bioinformatics*. 31 (8), 1296–1297. doi:10.1093/bioinformatics/btu817

Hullman, J., Drucker, S., Henry Riche, N., Lee, B., Fisher, D., and Adar, E. (2013). A deeper understanding of sequence in narrative visualization. *IEEE Trans. Vis. Comput. Graph* 19 (12), 2406–2415. doi:10.1109/TVCG.2013.119

Human abilities (2020). Disclaimer: online content. Available at: https://www.human-abilities.de/legal-notice/disclaimer/index.html.

Inequaligram (2020). Available at: http://inequaligram.net/.

ISO (2020). Available at: https://www.iso.org/obp/ui/#iso:std:iso, 9241:-11:ed-2:v1:en.

Jang, A., MacLean, D. L., and Heer, J. (2014). BodyDiagrams: improving communication of pain symptoms through drawing. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CA 94305 USA. Available at: https://idl.cs.washington.edu/files/2014-BodyDiagrams-CHI.pdf. doi:10.1145/2556288.2557223

Kandel, S., Parikh, R., Paepcke, A., Hellerstein, J. M., and Heer, J. (2012). Profiler: integrated statistical analysis and visualization for data quality assessment. Proceedings of the International Working Conference on Advanced Visual AVI '12, May 21-25. Available at: http://vis.stanford.edu/files/2012-Profiler-AVI.pdf.

Katoh, K., Rozewicki, J., and Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* 20 (4), 1160–1166. doi:10.1093/bib/bbx108

Kay, A., and Goldberg, A. (1977). Personal dynamic media. *Computer*, 10. 31–41. doi:10.1109/c-m.1977.217672

Kay, M., and Heer, J. (2015). *Beyond weber's law: a second look at ranking visualizations of correlation*. IEEE. doi:10.1109/TVCG.2015.2467671

Keim, D., Qu, H., and Ma, K. L. (2013). Big-data visualization. *IEEE Comput. Graph Appl.* 33 (4), 20–21. doi:10.1109/mcg.2013.54

Khoury, C. K., Kisel, Y., Kantar, M., Barber, E., Klirs, C., Ricciardi, V., et al. (2019). Science–graphic art partnerships to increase research impact. *Commun. Biol.* 2, 295. doi:10.1038/s42003-019-0516-1

Kim, Y., Correll, M., and Heer, J. (2019). Designing animated transitions to convey aggregate operations. *Computer Graphics Forum (Proc. EuroVis)*. 38 (3), 541–551. doi:10.1111/cgf.13709

Klement, M., and Walat, W. (2015). Multimediality and interactivity in e-learning. 1st international conference on lifelong learning and leadership for all. *At: Olomouc* Volume, 1.

Klipfolio (2020). What is a data dashboard? Available at: https://www.klipfolio.com/resources/articles/what-is-data-dashboard (Accessed April 2, 2021).

Knaffic, C. N. (2015). *Storytelling with data: a data visualization guide for business professionals*. Nov: John Wiley & Sons.

Kosara, R., Bendix, F., and Hauser, H. (2006). *Parallel sets: interactive exploration and visual analysis of categorical data transactions on visualization and computer graphics*, 12 (4). 558–568. doi:10.1109/TVCG.2006.76

Kosara, R., and Mackinlay, J. (2013). Storytelling: the next step for visualization. *Computer* 46 (5), 44–50. doi:10.1109/mc.2013.36

Kosara, R., Sahling, G. N., and Hauser, H. (2004). *Linking scientific and information visualization with interactive 3D scatterplots*, WSCG2004.

Kraak, M. J., and Ormeling, F. (2020). *Cartography: visualization of geospatial data*. London. Prentice-Hall.

Laevaert, C., and Le Goff, B. (2020). NTTA 2017 agenda. ESS vision 2020. Eurostat, sharing visualisation tools between national statistical institutes: a successful experience. Available at: https://ec.europa.eu/eurostat/cros/ntts2017programme/data/sessions/session_22.html.

Larson, K., and Czerwinksi, M. (1998). Web page design: implications of memory, structure and scent for information retrieval, *CHI '98 human factors in computing systems | april*.

Lee, B., Riche, N. H., Isenberg, P., and Carpendale, S. (2015). More than telling a story: transforming data into visually shared stories. *IEEE Comput. Graph Appl.* 35 (5), 84–90. doi:10.1109/MCG.2015.99

Lindt, I., Ohlenburg, J., Pankoke-Babatz, U., and Ghellal, S. (2007). A report on the crossmedia game epidemic menace. *Comput. Entertain.* 5 (1), 8. doi:10.1145/1219124.1219132

Lindt, I., Ohlenburg, J., Pankoke-Babatz, U., Ghellal, S., Oppermann, L., and Adams, M. (2005). *Designing cross media games. Pervasive'05, May*.

Liu, J., Xia, M., Dai, Z., Wang, X., Liao, X., Bi, Y., et al. (2017). Intrinsic brain hub connectivity underlies individual differences in spatial working memory. *Cereb. Cortex* 27 (12), 5496. doi:10.1093/cercor/bhw317

Lupton, E. (2017). *Design is storytelling*. 1 edition. Cooper Hewitt, Smithsonian Design Museum.

Manovich, L. (2011). what is visualizationVisual Studies. *Locative media and the digital visualisation of space, place and information*, Vol. 26.

Manovich, L. (2016). Social media inequality: definition, measurements, and application. Available at: http://manovich.net/index.php/projects/social-media-inequality.

Manovich, L. (2013). *Software takes command*. New York, United States: Bloomsbury Publishing.

Manovich, L. (2001). *The language of new media*. MIT Press.

Marcus, E. (2010). 2010: a publishing odyssey. *Cell* 140, 9. doi:10.1016/j.cell.2009.12.048

May, R. M. (2019). *Stability and complexity in model ecosystems.*
McCann, A. (2020). *Dueling Data*. Available at: https://duelingdata.blogspot.com/2017/12/3-types-of-data-visualizations.html.
Microsoft Visio (2020). Process diagrams in Visio. Available at: https://products.office.com/en-us/visio/flowchart-software.
Moere, A. V., and Purchase, H. (2011). On the role of design in information visualization. *Inf. Visualization* 10 (4), 356–371. doi:10.1177/1473871611415996
Mrvar, A., and Batagelj, V. (2016). Analysis and visualization of large networks with program package Pajek. *Casmodeling. Complex Adaptive Syst. Model.* 4. doi:10.1186/s40294-016-0017-8
Munzner, T. (2014). *Visualization analysis and design*. Florida, United States: CRC Press.
Nielsen Norman Group (2012). Usability 101: introduction to usability. Available at: https://www.nngroup.com/articles/usability-101-introduction-to-usability/.
Online Journalism Awards (OJAs) (2019). Available at: https://awards.journalists.org/.
Periscopic (2020). Available at: http//www.periscopic.com (Accessed April 2, 2021).
Popper, K. (1959). *The logic of scientific discovery (2002 pbk*. 2005 ebook ed. London, United Kingdom: Routledge. 978-0-415-27844-7.
Prakash, T. G. (2016). Data visualization and communication by big data. Int. Res. J. Eng. Technology. *(Irjet)* 03. 2395–0056. Available at: www.irjet.net.
Processing (2020). *Website: processing.*
Reas, C., and Fry, B. (2007). *Processing: a programming handbook for visual designers and artists*. MIT Press.
Riazi, S., and Norris, B. (2016). GraphFlow: workflow-based big graph processing. Conference: 2016 IEEE International Conference on Big Data. Washington, United States: IEEE. doi:10.1109/BigData.2016.7840993
Rothwell, D. J. (2004). *The company of others: an introduction to communication*. New York: McGraw-Hill.
Ruys, F. (2020). Infographic taxonomy. Available at: http://www.vizualism.nl/infographic-taxonomy/
SAP (2020). Lumira. Available at: https://www.sap.com/products/lumira.html.
SAP (2011). Navigable visualization of a hierarchical data structure. Available at: https://patents.google.com/patent/US9229609B2/en.
Satyanarayan, A., Moritz, D., Wongsuphasawat, K., and Heer, J. (2017). Vega-lite: a grammar of interactive graphics. *IEEE Trans. Vis. Comput. Graph* 23 (1), 341–350. doi:10.1109/TVCG.2016.2599030
Satyanarayan, A., and Heer, J. (2010). Lyra: an interactive visualization design environment. *Computer Graphics Forum* 33 (3), 351–360.
Schauer, S (2017). Once upon A time – 3 storytelling techniques for graphic design. Available at: https://bsodigital.com.au/upon-time-3-storytelling-techniquesgraphic-design/ (Accessed April 2, 2021).
Schneiderman, B. (1996). The eyes have it: a task by data type taxonomy for information visualizations. VL '96: Proceedings of the 1996 IEEE Symposium on Visual Languages.
Schramm, W. (1954). "How communication works", in *The process and effects of communication*. Editor W. Schramm (Urbana, Illinois: University of Illinois Press.), 3–26.
Scott, D. W. (2015). *Multivariate density estimation: theory, practice, and visualization.* John Wiley & Sons.
Segel, E., and Heer, J. (2010). Narrative visualization: telling stories with data. *IEEE Trans. Vis. Comput. Graph*, 16, 1139–1148. doi:10.1109/TVCG.2010.179
Shannon, C. E., and Weaver, W. (1949). *The mathematical theory of communication*. Urbana, Illinois: University of Illinois Press.
Siemens (2020). SIMATIC WinCC V7. Available at: https://new.siemens.com/.
Sofware, Uncharted. (2020). Uncharted products Available at: https://uncharted.software/product/.
SystemsBiology (2016). Computer-simulation of biological systems. YouTube. Available at: https://www.youtube.com/watch?v=dchHehGDfkc (Accessed April 2, 2021).
Terzidis, K. (2009). *Algorithms for visual design using the processing language*. Wiley.
Tufte, E. (1983). *The visual display of quantitative information*, 2. Cheshire, CT: Graphics Press., 9.
Uncharted Software (2018). System and method for data visualization using a synchronous display of sequential time data and on-map planning. Patent. Available at: https://patents.google.com/patent/US7688322B2/en (Accessed April 2, 2021).
Vega-Lite (2020). Vega-Lite – a grammar of interactive Graphics. Available at: https://vega.github.io/vega-lite/ (Accessed April 2, 2021).
Viégas, F. B., and Wattenberg, M. (2006). Communication-minded Visualization: a call to action. *IBM Syst. J.* 45 (4). doi:10.1147/sj.454.0801
W3C (2019). Web content accessibility guidelines (WCAG) 2.0. Available at: https://www.w3.org/TR/WCAG20/.
Walz, S. P. (2010). *Toward a ludic architecture: the space of play and games*. Architectural design: Lulu.com.
Ware, C. (2019). in *Information visualization: perception for design Morgan Kaufmann* (Elsevier).
Wattenberg, M., Viégas, F., and Johnson, I. (2016). *How to use t-SNE effectively*. Distill.
Weber, W., and Rall, H-M. (2016). Design and journalism – challenges and opportunities. A dialogue between two cultures. ICDHS. Available at: https://core.ac.uk/download/pdf/148757126.pdf. doi:10.5151/despro-icdhs2016-04_002
Weblyzard (2020). Cross media analysis. WebLyzard technology. Available at: https://www.weblyzard.com/cross-media-analysis/ (Accessed April 2, 2021).
Wilkinson, L. (1999). *The grammar of graphics*. Springer.
Willett, W., Heer, J., and Agrawala, M. (2012). Strategies for crowdsourcing social data analysis. in *ACM human factors in computing systems*. New York, United States: ACM. 227–236.
Wojdynski, B. W., and Kalyanaraman, S. (2015). The three dimensions of website navigability: explication and effects. *J. Assn Inf. Sci. Tec*, 67. Wiley Online Library., 454. doi:10.1002/asi.23395
Wolfram, S. (2018). *Cellular automata and complexity: collected papers*. Florida, United States: CRC Press.
Wongsuphasawat, K., Qu, Z., Moritz, D., Chang, R., Ouk, F., Anand, A., et al. (2017). Voyager 2: augmenting visual analysis with partial view specifications. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems.
Wu, T., Ribeiro, M. T., Heer, J., and Weld, D. S. (2019). Errudite: scalable, reproducible, and testable error analysis. Proceedings of the 57th Conference of the Association for Computational. doi:10.18653/v1/p19-1073
Xu, G., Wu, Y., Minshall, T., and Zhou, Y. (2017). Exploring innovation ecosystems across science, technology, and business: A case of 3D printing in China. *Technol. Forecast. Soc. Change* 136. doi:10.1016/j.techfore.2017.06.030
Yang, W. J. (2018). *Handbook of flow visualization*. New York, United States: Routledge. doi:10.1201/9780203752876
Yau, N. (2013). *Data points: visualization that means something*. Wiley.
Yosinski, J., Clune, J., Nguyen, A., Fuchs, Th., and Lipson, H. (2015). Understanding neural networks through deep visualization. *Deep learning workshop*, 31. San Diego, United States: International Conference on Machine Learning.
Zimmermann, R. (2011). Stories we have lived, Stories we have learned. *About Early Childhood Development Programmes The Effectiveness Initiative. Bernard van Leer Foundation. Complete quote: "The exchange of experiences, supported by good example, can become an efficient motivational strategy in encouraging participation in common efforts".*
ZoomChart (2020). ZoomChart network for power BI. Available at: https://zoomcharts.com/introducing–network-chart-custom-visual-for-microsoft-power-bi-mimg-1109-mega (Accessed April 2, 2021).
Zuo, R., Carranza, E. J. M., and Wang, J. (2016). Spatial analysis and visualization of exploration geochemical data. *Earth-Science Rev.* 158, 9–18. doi:10.1016/j.earscirev.2016.04.006

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Research Metrics and Analytics

ORIGINAL RESEARCH
published: 22 April 2021
doi: 10.3389/frma.2021.651991

Highly Cited Papers at the Spanish Domestic Level

Carlos García-Zorita[1,2], *Sergio Marugán*[1], *Daniela De Filippo*[1,2]* *and Elías Sanz-Casado*[1,2]

[1] Department of Library and Information Science, Carlos III University of Madrid (UC3M), Getafe, Spain, [2] INAECU Research Institute for Higher Education and Science Autonomous University of Madrid (UAM) Carlos III University of Madrid (UC3M), Madrid, Spain

This paper presents a methodological proposal based on the identification of highly cited papers (HCPs) at domestic-level in the Spanish Public University System (SUPE), in order to find the most outstanding publications in the local context. The principal aim is to detect different activity and impact profiles among Spanish universities and differentiate those institutions that play a more significant role. To determine which and how many are the highly cited papers at the domestic level (HCP-DL) collected in the Web of Science, three citation thresholds (1, 5, and 10%) were established. Thematic classification in Incites/Essential Science Indicators areas is used. The results show a preponderance of HCPs in the field of Space Science, while the polytechnic universities have high visibility in the Computer Science area. It has been observed that the presence of HCPs in a given area is involved with universities specialized in teaching and research activities. In absolute terms, the big non-specialized universities are major producers of HCPs and hold the leading positions in our results. However, when efficiency is analyzed in relative terms, some small, specialized universities reveal themselves to be more efficient at producing HCPs (% of HCPs or citations per HCP). We think that this methodology, due to its simplicity, its ease of calculation, and the knowledge it provides, can be very useful to analyze the national systems of any country, in order to know the impact and visibility of the research carried out in its scientific institutions or research areas.

Keywords: highly cited papers, Spanish universities, visibility indicators, impact indicators, higher education institutions

OPEN ACCESS

Edited by:
Juan Ignacio Gorraiz,
Universität Wien, Austria

Reviewed by:
Aparna Basu,
National Institute of Science, Technology and Development Studies (CSIR), India
Jamal El Ouahi,
Leiden University, Netherlands

***Correspondence:**
Daniela De Filippo
dfilippo@bib.uc3m.es

Specialty section:
This article was submitted to Research Assessment, a section of the journal Frontiers in Research Metrics and Analytics

Received: *11 January 2021*
Accepted: *26 March 2021*
Published: *22 April 2021*

Citation:
García-Zorita C, Marugán S, De Filippo D and Sanz-Casado E (2021) Highly Cited Papers at the Spanish Domestic Level. Front. Res. Metr. Anal. 6:651991. doi: 10.3389/frma.2021.651991

INTRODUCTION

Right from the creation of the impact factor (IF) by Garfield (1955), the field of scientometrics assumed this indicator as a measure of analysis of scientific performance based on the number of citations a journal achieves. Thus, the impact of an author or an institution was held to be equal to the impact of the journals where their papers were published.

Although several studies have shown that journal's IF does not accurately reflect the impact of each individual article (Seglen, 1997; Garfield, 2006), many institutions and national evaluation systems still use it.

The debate over the use (and abuse) of the IF, and criticism of its application and reformulations, could be reduced if, as some authors explain, it is assumed that the IF has played a meritorious role in identifying influential journals and should continue to be used as an indicator of competitiveness and reputation. That is, as an indicator of the capacity of an author or institution to publish in journals with a high publication demand (Orduña-Malea et al., 2016).

With the appearance of Web 2.0. there has been an unprecedented change in the world of scientific activity production and dissemination. The use of new platforms to generate and share data and research results as well as the creation of digital identities have influenced the field of research evaluation (De Filippo and Sanz-Casado, 2018). Traditional studies based on bibliometrics can be complemented with new indicators such as altmetrics, which measure the interest that research arouses in society and have had a particular impact since their appearance in 2010 (Priem and Hemminger, 2010). One of the main advantages of these indicators is that, since the data are presented at the article level, a study's impact can be evaluated without considering the quality or visibility of the journal of publication (Neylon and Wu, 2009). As Martín-Martín et al. (2018) comment, "Since the classic study by Bollen et al. (2009), where the data came primarily from usage logs provided by publishers, many papers have been published on the nature of online article-level metrics." Some of these studies have tried to correlate traditional citation with citation over different platforms that offer indicators of social media impact.

The possibility of analysis provided by new social media and platforms has led some authors (Orduña-Malea et al., 2016; Martín-Martín et al., 2018) to mention the emergence of a new line of bibliometric research, ALMetrics (author level metrics), which analyses the performance of authors by measuring all the dimensions of their intellectual activity. Without a doubt, many challenges arise with these options for assessment at the document and author level, not only from a technical point of view but also, and especially, for the study and evolution of impact and visibility.

From the point of view of research evaluation, both at individual and institutional level, another indicator that has started to be used in the last decades is highly cited papers (HCPs). One reasons for this is the increasing focus on scientific excellence in scientific policy (van-Raan, 2000). Science policy is increasingly interested in scientific excellence given its new public management tools (Aknes, 2003; Lamont, 2012). "Many countries are moving toward research policies that emphasize excellence; consequently; they develop evaluation systems to identify universities, research groups, and researchers that can be said to be "excellent" (Danell, 2011). This was shown in a diverse studies as a benchmarking study from the European Commission in which HCPs were used as indicators for comparing the research performance of the EU countries (European Commission, 2001). Highly cited papers have also been applied as indicators in case studies of research groups and some authors concluded that highly cited research papers do represent useful indicators for identifying "worldclass" research (Tijssen et al., 2002).

In recent years, the use of highly cited articles has become increasingly common and indicators, such as those developed by Clarivate Analytics (2020), are being widely used for institutional evaluation. In the field of higher education, indicators of excellence in research have also been developed, such as those offered by the Ranking of SCImago Institutions (Bornmann et al., 2012), the mapping of excellence (Bornmann et al., 2014), and the Ranking Leiden (Waltman et al., 2012). Despite the increasingly frequent use of these indicators, these indices are not exempt from criticism, both from the methodological point of view and their application (Hu et al., 2018), so it is essential to continue developing the research in this field.

In this line, some authors also mention that it is urgent to look further into the phenomenon of HCPs, especially in small and peripheral countries, where the need to be selective is largest, the citation indicator is more uncertain than in core countries (Aknes, 2003).

In this context, highly cited articles have been considered as potential candidates for identifying and monitoring "excellent" scientific research. A wide range of options lies open for the analysis of the scientific activity of institutions such as universities, one of the main producers of knowledge, whose evaluation requires precise tools.

Institutional-Level Metrics: Evaluation in Universities

The analysis and evaluation of the research activity carried out in the institutions has been a decisive step to really know the scope of these activities, make proposals, and offer society the necessary transparency of its efficient management of the resources allocated to the research carried out in these institutions. In this way, Szomszor et al. (2021) state that "Research evaluation may be seen as a reflection of a broader societal shift to institutional managerialism and public sector accountability." However, it is within higher education institutions where evaluation has been more ingrained and where it is playing a more decisive role. The reasons why this effects have occurred are several, for example, accountability to society for the activities they carry out, the proper management of the financial resources they receive, or knowing how the scientific productivity of their academic staff evolves. One of the countries that first considered the need to evaluate its higher education institutions was the United Kingdom, where the first national Research Selectivity Exercise was introduced in 1986 and led to a more formalized and structured Research Assessment Exercise (RAE) from 1992 (Szomszor et al., 2021). This evaluation process has currently changed its name to Research Excellence Framework (REF) (REF, 2020) and it has had multiple counterparts in different countries (Sanz-Casado et al., 2013), especially in the Nordic countries (Sivertsen, 2018) and in Australia where the Australian Research Council (ARC) conducted the first Excellence in Research for Australia (ERA) evaluation in 2010 (ARC, 2019). These institutional evaluation processes have gained renewed importance with the emergence of international university rankings since 2003.

Rankings such as the Shanghai (ARWU), Times Higher Education (THE), and QS have had great impact, and they have served to provide information on higher education institutions around the world. These rankings, which have spurred the debate about the quality and performance of higher education systems, have had a considerable impact on our global society in light of the internationalization of higher education. That, in turn, has heightened global competition and induced proliferation of this type of studies (De Filippo et al., 2012). However, criticism of their methodology and implementation has also been plentiful (Liu and Cheng, 2005; van-Raan, 2005; Buesa et al., 2009). The methodology used to formulate these rankings cannot deliver reliable data for more than 700–1,200 institutions, in the light of the wide range of variation in the non-top-ranked universities. Another frequent criticism is that this methodology may therefore be regarded as "elitist," inasmuch as it entails excluding the vast majority of the world's universities (De Filippo et al., 2012).

The need to complement the information provided by international classifications has fostered the development of some initiatives with data at the national level. Several rankings have been developed in Spain, such as the Multidimensional Index of University Quality (Buesa et al., 2009), the Research Ranking of Spanish Public Universities (Buela-Casal et al., 2011), the General and Area Ranking of Spanish University Institutions

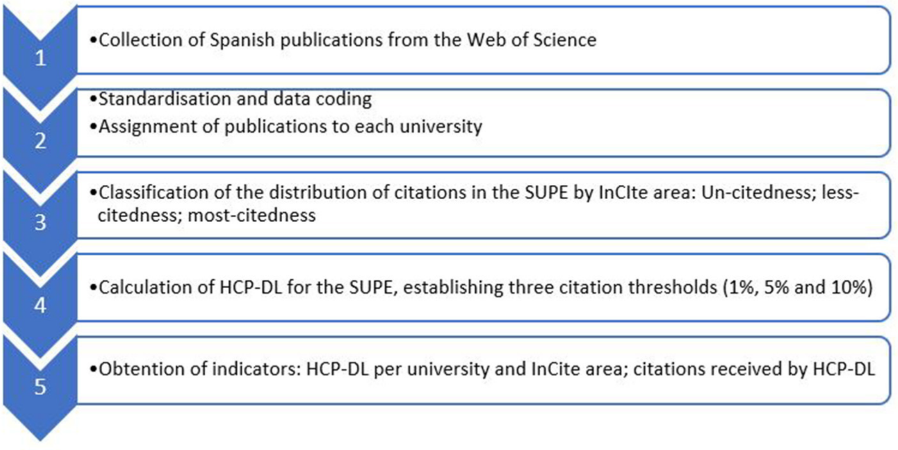

FIGURE 1 | Methodological steps.

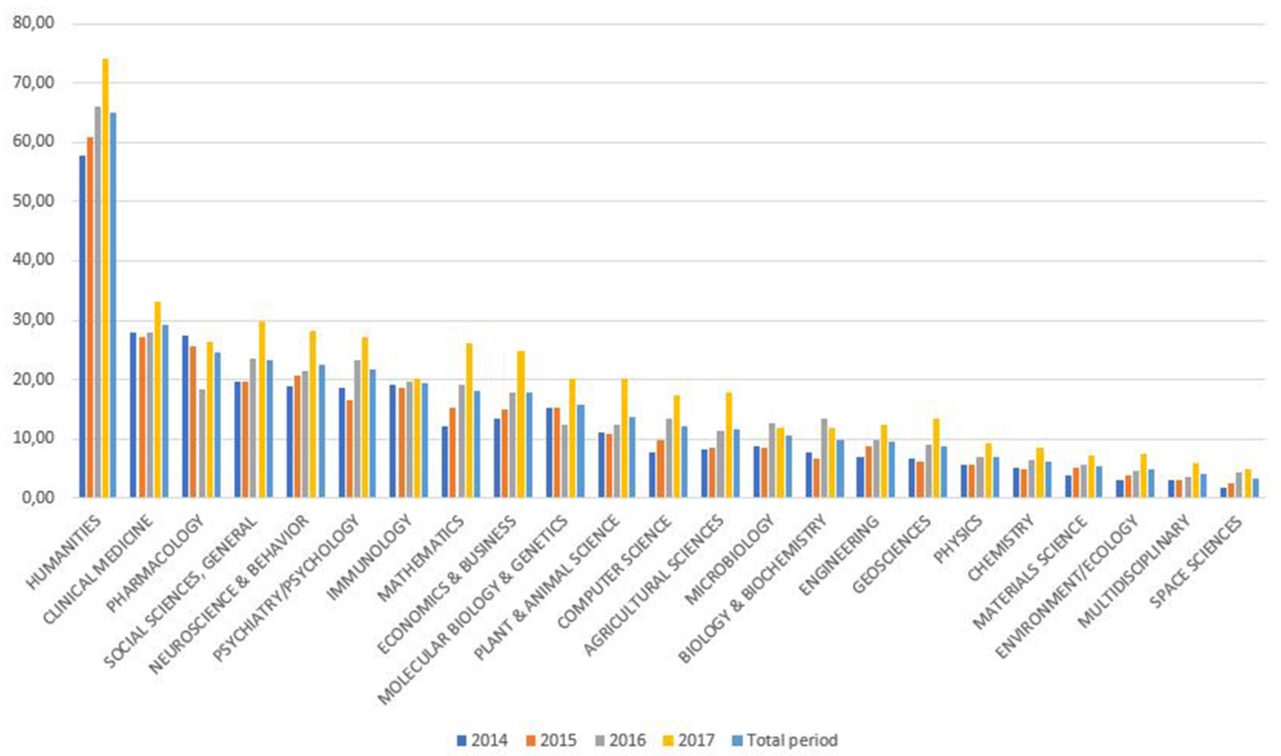

FIGURE 2 | Percentage of un-citedness in SUPE by Incites/Essential Science Indicators area.

(Corera et al., 2010), the I-UGR Ranking of Spanish Universities (Torres-Salinas et al., 2011), and the Observatory of Research Activity of the Spanish University (IUNE) with annual updates from 2012 (Sanz-Casado et al., 2011, 2013; De Filippo et al., 2014).

The IUNE Observatory, created by the 4 Universities Alliance (A4U), has the support of the Spanish Ministry of Universities and offers aggregate information on seven dimensions (teaching staff, scientific recognition, scientific activity, innovation, research training capacity, competitiveness, and funding). The data are obtained from official and public sources and are presented through 48 indicators. The scientific publications are collected from the Web of Science core collection (www.iune.es).

One of IUNE's basic premises is the presentation of a wide range of indicators to present a simple, transparent picture of each institution's scientific activity, trying to account for the variety of profiles in existence. This is possible because IUNE considers a large number of indicators related to the scientific and knowledge transfer activity of Spanish universities, unlike other rankings that assign the greatest weight to bibliometric indicators.

The great variety of data obtained enables basic information to be displayed over the web by university, by major fields of knowledge and in terms of the university system as a whole (Bautista-Puig et al., 2020). Among the indicators within the IUNE framework, different metrics are being developed that consider the document as the object of study. Some of them, related to impact, are presented below. The calculation of indicators at the local (country) level is key to making comparisons between institutions in the same context. In this paper we present the methodology developed to calculate HCPs without using international comparison that may be far removed from local practices.

OBJECTIVES

The research presented in this paper has been aimed at the following objectives:

- To develop a methodological proposal based on the identification of HCPs in domestic systems, such as the Spanish Public University System (SUPE), in order to

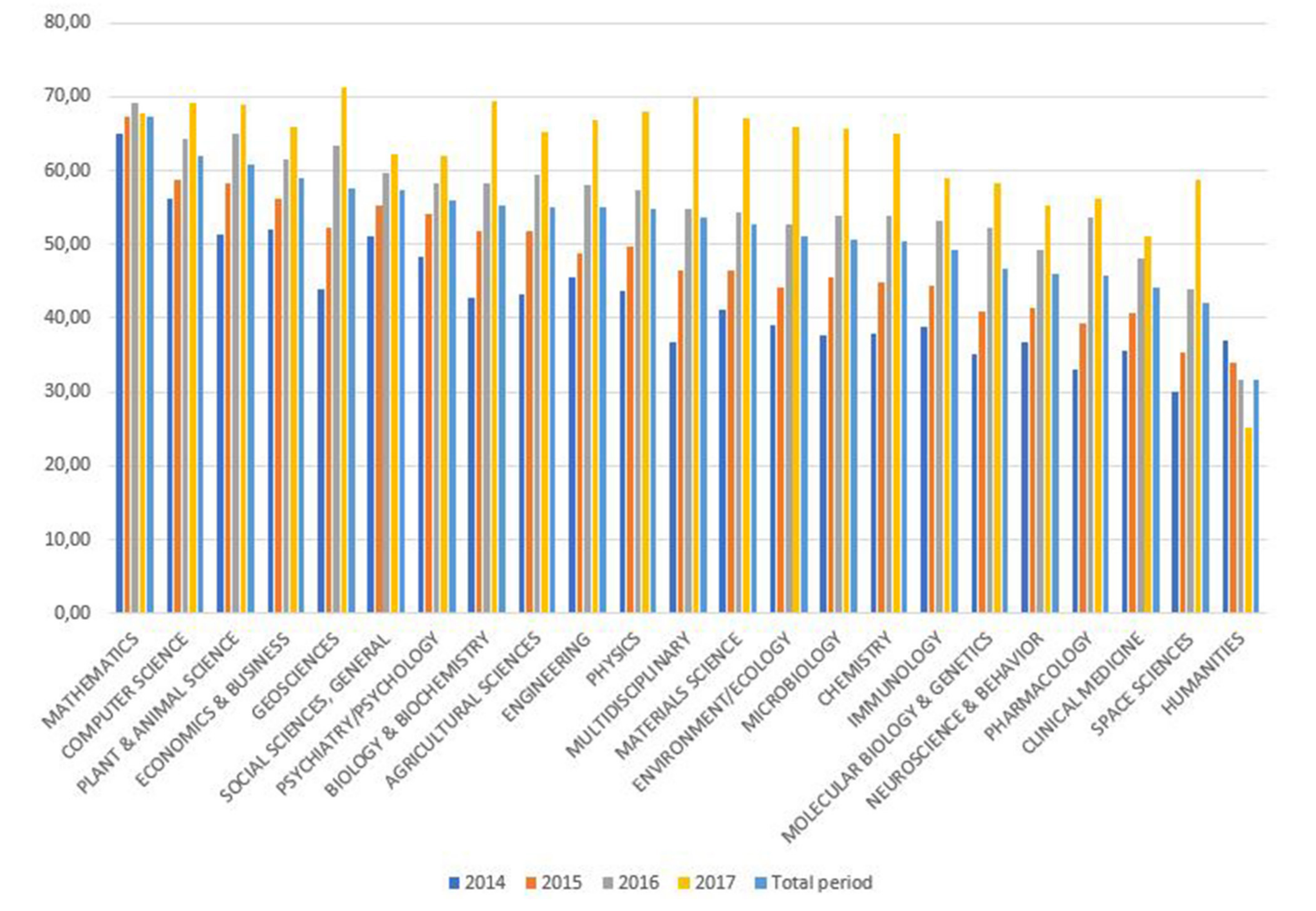

FIGURE 3 | Percentage of less-citedness in SUPE by Incites/Essential Science Indicators area.

find the most outstanding publications in the Spanish scientific context.
- To detect different activity and impact profiles among Spanish universities. This will make it possible to differentiate those universities that play a more significant role in these two important aspects.

METHODOLOGY

This study uses a specific methodology to explore each Spanish public university's highly cited papers at the domestic level (HCP-DL), calculated in relation to the total scientific production of the SUPE in the Web of Science citation indexes.

The data obtained from the IUNE Observatory, which includes publications collected from the three main databases of the Web of Science core collection (Science Citation Index, Social Science Citation Index, Arts and Humanities Citation Index), are used as a source of information.

To identify the SUPE's production, a system based on regular expressions is used to encode and normalize the signature of each document. Regular expressions are patterns used to find a certain combination of characters within a text string (Ruslan, 2003). This enables each university's publications to be identified by searching for different signature variants in the "address" field. This system assigns publications to each institution using the total count of documents (one publication is counted for each signatory institution). Although there are standardization options such as the "enhanced organization" of Web of Science, the identification by regular expressions, which has been used at IUNE for more than a decade, has different advantages. On the one hand, it allows a "strict" attribution of documents, i.e., it only considers the university's own production (not including documents produced by university hospitals, health centers, consortia, etc., in which the explicit signature of a university does not appear). It also allows information to be retrieved from incomplete signatures (only postal addresses, names of centers or departments, which clearly belong to a university). With this system, some universities see an increase in their output compared to the direct WoS query (as greater flexibility and breadth in the identification of university documents is possible), while others see a reduction in their output (by eliminating documents considered to be "university documents" but without an explicit signature). APPENDIX I (of the **Supplementary Material**) provides a comparative table retrieving information from WoS and IUNE.

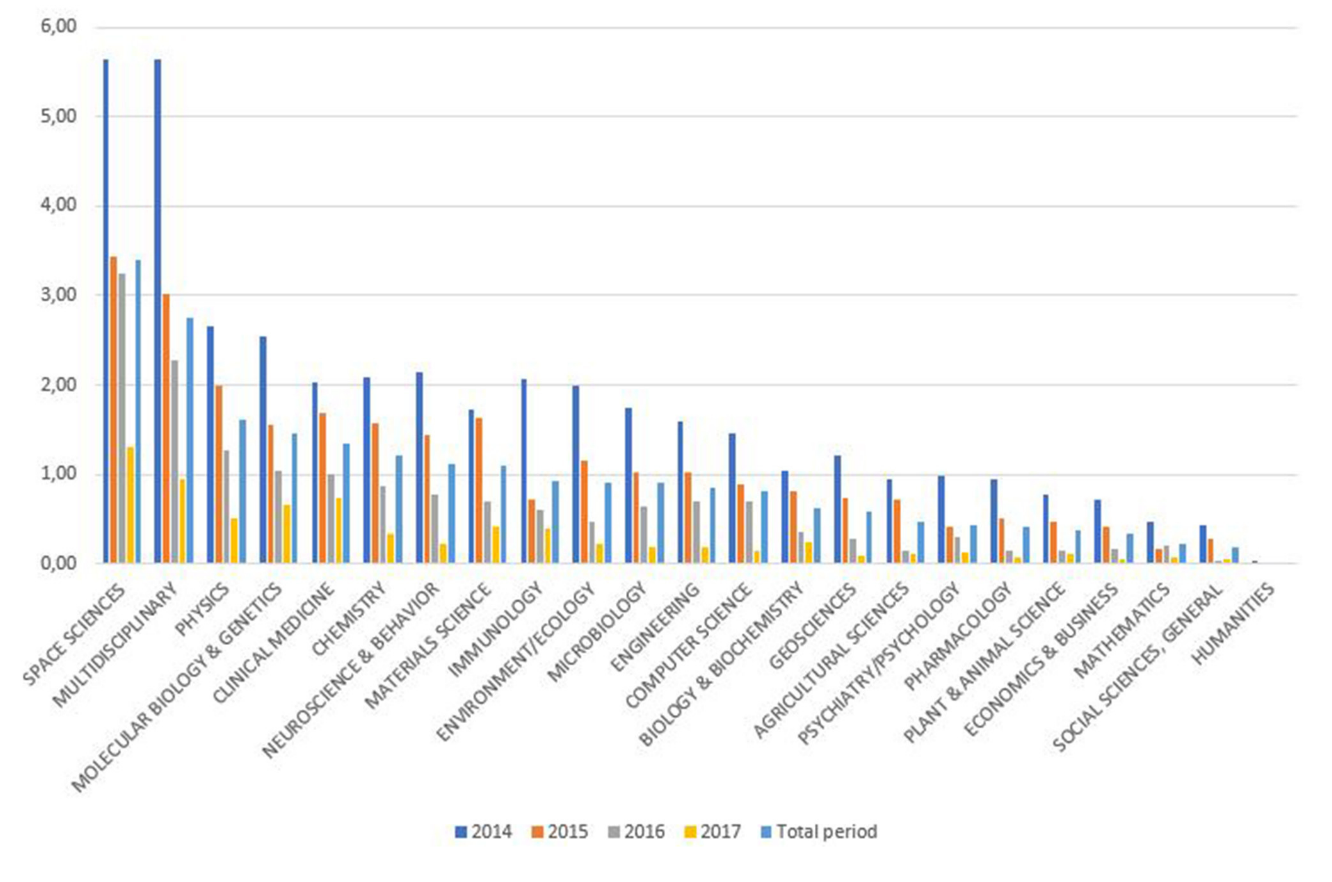

FIGURE 4 | Percentage of most-citedness in SUPE by Incites/Essential Science Indicators area.

The use of regular expressions for the identification of the production of one university (Universidad Carlos III de Madrid) is also presented.

The Spanish university system currently has 83 universities (50 public and 33 private). This study analyses the production of public universities since they have more-intensive research activity and produce more than 95% of the Spanish university system's publications (Casani et al., 2013). The list of public universities and their acronyms is shown in IUNE Glosario (IUNE, 2020).

The study is carried out per year, and the analysis period includes the publications of 2014–2017 (citations were collected in February 2020). This citation window has been chosen so that publications have at least 2 years of citations; otherwise, the data may be very distorted. The study has been carried out considering the production by subject area, given that there are differences in the dynamics of production, impact, and visibility of the different scientific disciplines (Aknes, 2003; Aksnes and Sivertsen, 2004). For this purpose, the thematic aggregation carried out by Web of Science (Incites/Essential Science Indicators areas) was considered. Twenty-two areas were considered, plus one more area, humanities, to differentiate the production of this field from that of Social Sciences, which has important differences (Huang and Chang, 2008). Lists were thus produced for each year and area, with the publications ordered by the number of citations received. This process allows us to determine the minimum number of citations that a publication must have obtained in order to be considered a HCP.

The first relevant information to find is the distribution of citations by thematic area in the SUPE. The publications were classified into different groups according to their impact. There are various definitions of what counts as a highly cited article. Basically two different approaches can be identified, involving absolute, or relative thresholds (Aknes, 2003). Therefore, in this study the publications have been classified into three groups: (i) Un-citedness (documents without citations up to the time of data collection); (ii) less-citedness (documents receiving between 1 and 10 citations); (iii) most-citedness (documents receiving more than 100 citations). These limits were established for convenience and for simplicity of comprehension. This calculation provides information on the general dynamics of citation in the SUPE.

Next, the first step for calculating HCP-DL is to determine which and how many are the HCP-DL that are collected in the Web of Science, establishing three citation thresholds (1, 5, and 10%). Once the papers are ranked, the number of citations that a publication needs to be considered HCP-DL is selected.

Some indicators are calculated from this data:

- number of HCP-DL at public Spanish universities and by InCite area.
- percentage of highly cited papers (HCP-DL) for each university.

TABLE 1 | HCP-DL lower limit, minimum number of citations needed to qualify as HCP-DL.

Incites/Essential science indicators area	1%				5%				10%			
	2014	2015	2016	2017	2014	2015	2016	2017	2014	2015	2016	2017
Agricultural Sciences	96	84	52	41	48	39	29	21	35	29	21	15
Biology and Biochemistry	101	92	67	48	51	41	30	22	36	27	20	15
Chemistry	157	129	89	55	62	53	40	27	43	36	28	19
Clinical Medicine	146	131	100	77	56	48	32	25	35	29	21	15
Computer Science	122	97	71	47	48	36	27	19	30	25	18	12
Economics and Business	94	70	49	36	49	33	26	15	33	24	17	10
Engineering	122	101	80	49	56	46	34	24	38	32	24	17
Environment/Ecology	152	111	81	54	62	53	41	27	43	39	29	20
Geosciences	115	91	67	38	47	40	30	19	34	27	20	13
Humanities	26	21	15	9	11	11	6	4	6	5	4	2
Immunology	150	84	82	53	58	44	34	26	40	28	22	18
Materials Science	122	125	85	60	60	52	38	29	40	37	28	20
Mathematics	63	50	33	25	28	22	17	12	19	15	12	8
Microbiology	132	103	78	59	58	44	34	26	39	32	23	17
Molecular Biology and Genetics	159	126	108	68	64	51	38	27	42	33	26	17
Multidisciplinary	326	175	170	98	115	68	59	31	62	44	36	20
Neuroscience and Behavior	132	143	85	47	59	46	37	22	40	30	24	14
Pharmacology	99	71	55	37	45	38	31	21	30	27	21	14
Physics	184	148	124	73	68	55	41	28	45	37	27	18
Plant and Animal Science	90	75	51	34	41	34	24	16	29	22	17	11
Psychiatry/Psychology	101	69	53	36	45	33	25	17	29	22	16	11
Social Sciences. General	73	60	44	29	38	29	21	14	25	21	15	9
Space Sciences	376	224	231	124	110	85	68	44	71	54	45	27

- average number of citations per document received by HCP-DL.

Figure 1 shows the methodological steps followed.

RESULTS

From 2014 to 2017 the SUPE published 218,779 documents in the Web of Science core collection. The main results of each phase of the HCP-DL calculation process are presented below.

Citedness Distribution

Figures 2–4 show the distribution of citedness on three levels. The data is sorted in descending order by the total number of documents for each subject area in the period. Tables with the corresponding percentage values are given in APPENDIX II (**Supplementary Material**). The percentage of documents not cited is presented in **Figure 2**, distributed by thematic area and year. Humanities is the field with the highest proportion of uncited documents (65%). Other areas, such as Clinical Medicine, Neurosciences, Pharmacology, Mathematics,

TABLE 2 | HCP-DLs by university and incites/essential science indicators area in all three top citation groups (absolute values).

University	AGRICULTURAL SCIENCES	BIOLOGY & BIOCHEMISTRY	CHEMISTRY	CLINICAL MEDICINE	COMPUTER SCIENCE	ECONOMICS & BUSINESS	ENGINEERING	ENVIRONMENT/ECOLOGY	GEOSCIENCES	HUMANITIES	IMMUNOLOGY	MATERIALS SCIENCE	MATHEMATICS	MICROBIOLOGY	MOLECULAR BIOLOGY	MULTIDISCIPLINARY	NEUROSCIENCE & BEHAVIOR	PHARMACOLOGY	PHYSICS	PLANT & ANIMAL SCIENCE	PSYCHIATRY/PSYCHOLOGY	SOCIAL SCIENCES. GENERAL	SPACE SCIENCES	Total Arts
HCP-DL (1%)																								
UB	12	8	21	130	5	1	17	21	7	13	17	8	4	7	34	14	27	11	27	17	35	15	9	399
UAB	5	8	22	83	3	6	29	19	13	7	8	14	2	8	16	9	24	7	44	18	14	21	15	326
UV	20	3	26	56	1	4	18	15	16	6	3	12	6	5	8	7	4	7	60	5	8	2	8	247
UAM	2	3	25	34	4	3	11	14	1	2	2	11	4	6	12	7	2	2	58	2	7	10	16	207
UGR	14	3	6	8	36	5	10	4	10	4	1	1	13	0	5	2	2	3	32	3	6	12	18	173
UCM	9	1	13	32	1	3	9	9	8	8	5	8	8	5	3	4	6	4	7	9	8	9	2	147
UPF	1	8	2	13	1	4	4	6	0	8	3	0	7	6	26	19	8	0	0	2	13	25	0	137
EHU	6	2	27	5	8	0	21	10	4	20	0	13	3	2	6	5	6	5	16	4	6	3	1	124
USC	4	4	14	13	2	4	7	2	2	0	0	1	8	2	7	6	1	5	20	0	2	3	1	93
UNIOVI	0	0	8	9	1	0	4	2	0	3	1	2	1	3	10	5	1	0	18	2	5	3	18	89
HCP-DL (5%)																								
UB	51	41	128	529	11	18	71	82	39	68	57	43	20	33	133	52	134	60	102	68	135	68	37	1594
UAB	20	35	106	413	19	38	95	90	54	67	48	53	29	49	66	43	104	42	208	88	58	92	51	1482
UAM	31	21	109	159	13	11	53	52	12	13	16	52	34	19	55	34	20	17	315	12	38	46	91	988
UV	78	18	114	197	6	24	63	53	40	33	21	42	17	29	70	24	19	19	261	15	34	44	41	961
UGR	59	14	51	75	110	23	40	31	50	31	6	10	51	8	29	8	17	17	154	12	31	70	65	780
UCM	24	17	87	151	9	16	53	32	52	37	27	38	41	27	23	23	36	38	49	35	38	35	22	718
EHU	28	9	184	45	23	5	121	55	20	80	3	85	15	7	32	34	23	21	100	22	17	37	8	683
UPF	4	28	6	103	13	24	15	32	4	34	21	0	34	27	122	66	33	7	4	10	41	103	0	586
US	25	8	60	45	21	31	82	22	8	24	14	15	47	24	38	6	11	17	15	22	11	29	0	435
USC	23	9	50	80	5	11	29	13	6	16	0	14	22	12	32	13	5	21	83	9	10	13	8	396
HCP-DL (10%)																								
UB	102	81	245	976	21	36	131	150	83	136	112	95	33	71	254	120	254	106	207	131	264	142	69	3030
UAB	48	74	182	793	45	60	180	153	101	113	106	99	58	94	139	81	173	76	350	167	119	175	114	2732
UAM	58	44	204	321	27	22	88	74	17	39	33	84	76	37	139	66	53	36	529	42	70	85	174	1839
UV	124	33	222	351	17	56	104	93	65	68	33	67	35	55	114	43	39	41	449	37	83	117	89	1773
UGR	102	34	114	183	177	43	94	68	90	57	11	31	99	18	68	28	39	33	240	28	76	148	110	1492
UCM	48	32	172	330	28	32	107	54	106	61	57	83	73	57	60	40	72	77	111	66	75	70	42	1409
EHU	49	22	350	116	46	22	213	94	36	117	4	158	43	19	62	55	51	44	212	45	60	84	16	1368
UPF	11	48	11	206	28	48	29	65	5	82	40	0	58	39	229	101	55	17	6	19	72	175	0	1076
US	54	13	125	100	52	52	167	50	16	49	27	38	98	46	75	14	20	27	53	46	25	51	0	879
UPV	67	11	207	26	81	38	186	51	13	22	1	43	82	10	26	26	6	9	69	94	6	41	5	840

Color shades indicates the magnitude (green are universities and subject areas with higher values and red with lower values).

Immunology, and Molecular Biology, also have high un-citedness percentages. On the other hand, Space Science only has 3.4% non-cited documents and is the area with the highest visibility.

The proportions of less-cited documents receiving between 1 and 10 citations are similar in most areas, with a percentage range of between 50 and 60% (**Figure 3**). The area with the lowest proportion of less-cited documents is Humanities (31.8%), and Mathematics has the highest average (67.3%).

Figure 4 shows the documents with more than 100 citations, which reach higher levels only in the areas of Space Science (3.4%) and, to a lesser extent, Multidisciplinary (2.76%).

Highly Cited Papers at the Domestic Level

This section presents the distribution of citations received by documents published by SUPE universities, establishing three dynamic thresholds that vary according to the year of publication and the Incites/Essential Science Indicators area in which the journal is classified.

In order to define the conditions a paper must fulfill to be considered an HCP-DL, a minimum citation threshold is established, by area and year, for the top 1, 5, and 10% of the most-cited documents. These limits are shown in **Table 1**, where, for example, an article published in 2014 in Space Sciences needs 376 citations (or more) in 2020 to place among the top 1% of the most-cited documents in its area, but it needs only 124 citations

TABLE 3 | HCP-DLs by university and incites/essential science indicators area in all three top citation groups (percentages).

University	AGRICULTURAL SCIENCES	BIOLOGY & BIOCHEMISTRY	CHEMISTRY	CLINICAL MEDICINE	COMPUTER SCIENCE	ECONOMICS & BUSINESS	ENGINEERING	ENVIRONMENT/ECOLOGY	GEOSCIENCES	HUMANITIES	IMMUNOLOGY	MATERIALS SCIENCE	MATHEMATICS	MICROBIOLOGY	MOLECULAR BIOLOGY	MULTIDISCIPLINARY	NEUROSCIENCE & BEHAVIOR	PHARMACOLOGY	PHYSICS	PLANT & ANIMAL SCIENCE	PSYCHIATRY/PSYCHOLOGY	SOCIAL SCIENCES, GENERAL	SPACE SCIENCES	Total Arts
HCP-DL (1%)																								
UB	3.0%	2.0%	5.3%	32.6%	1.3%	0.3%	4.3%	5.3%	1.8%	3.3%	4.3%	2.0%	1.0%	1.8%	8.5%	3.5%	6.8%	2.8%	6.8%	4.3%	8.8%	3.8%	2.3%	399
UAB	1.5%	2.5%	6.7%	25.5%	0.9%	1.8%	8.9%	5.8%	4.0%	2.1%	2.5%	4.3%	0.6%	2.5%	4.9%	2.8%	7.4%	2.1%	13.5%	5.5%	4.3%	6.4%	4.6%	326
UV	8.1%	1.2%	10.5%	22.7%	0.4%	1.6%	7.3%	6.1%	6.5%	2.4%	1.2%	4.9%	2.4%	2.0%	3.2%	2.8%	1.6%	2.8%	24.3%	2.0%	3.2%	0.8%	3.2%	247
UAM	1.0%	1.4%	12.1%	16.4%	1.9%	1.4%	5.3%	6.8%	0.5%	1.0%	1.0%	5.3%	1.9%	2.9%	5.8%	3.4%	1.0%	1.0%	28.0%	1.0%	3.4%	4.8%	7.7%	207
UGR	8.1%	1.7%	3.5%	4.6%	20.8%	2.9%	5.8%	2.3%	5.8%	2.3%	0.6%	0.6%	7.5%	0.0%	2.9%	1.2%	1.2%	1.7%	18.5%	1.7%	3.5%	6.9%	10.4%	173
UCM	6.1%	0.7%	8.8%	21.8%	0.7%	2.0%	6.1%	6.1%	5.4%	5.4%	3.4%	5.4%	5.4%	3.4%	2.0%	2.7%	4.1%	2.7%	4.8%	6.1%	5.4%	6.1%	1.4%	147
UPF	0.7%	5.8%	1.5%	9.5%	0.7%	2.9%	2.9%	4.4%	0.0%	5.8%	2.2%	0.0%	5.1%	4.4%	19.0%	13.9%	5.8%	0.0%	0.0%	1.5%	9.5%	18.2%	0.0%	137
EHU	4.8%	1.6%	21.8%	4.0%	6.5%	0.0%	16.9%	8.1%	3.2%	16.1%	0.0%	10.5%	2.4%	1.6%	4.8%	4.0%	4.8%	4.0%	12.9%	3.2%	4.8%	2.4%	0.8%	124
USC	4.3%	4.3%	15.1%	14.0%	2.2%	4.3%	7.5%	2.2%	2.2%	0.0%	0.0%	1.1%	8.6%	2.2%	7.5%	6.5%	1.1%	5.4%	21.5%	0.0%	2.2%	3.2%	1.1%	93
UNIOVI	0.0%	0.0%	9.0%	10.1%	1.1%	0.0%	4.5%	2.2%	0.0%	3.4%	1.1%	2.2%	1.1%	3.4%	11.2%	5.6%	1.1%	0.0%	20.2%	2.2%	5.6%	3.4%	20.2%	89
HCP-DL (5%)																								
UB	3.2%	2.6%	8.0%	33.2%	0.7%	1.1%	4.5%	5.1%	2.4%	4.3%	3.6%	2.7%	1.3%	2.1%	8.3%	3.3%	8.4%	3.8%	6.4%	4.3%	8.5%	4.3%	2.3%	1594
UAB	1.3%	2.4%	7.2%	27.9%	1.3%	2.6%	6.4%	6.1%	3.6%	4.5%	3.2%	3.6%	2.0%	3.3%	4.5%	2.9%	7.0%	2.8%	14.0%	5.9%	3.9%	6.2%	3.4%	1482
UAM	3.1%	2.1%	11.0%	16.1%	1.3%	1.1%	5.4%	5.3%	1.2%	1.3%	1.6%	5.3%	3.4%	1.9%	5.6%	3.4%	2.0%	1.7%	31.9%	1.2%	3.8%	4.7%	9.2%	988
UV	8.1%	1.9%	11.9%	20.5%	0.6%	2.5%	6.6%	5.5%	4.2%	3.4%	2.2%	4.4%	1.8%	3.0%	7.3%	2.5%	2.0%	2.0%	27.2%	1.6%	3.5%	4.6%	4.3%	961
UGR	7.6%	1.8%	6.5%	9.6%	14.1%	2.9%	5.1%	4.0%	6.4%	4.0%	0.8%	1.3%	6.5%	1.0%	3.7%	1.0%	2.2%	2.2%	19.7%	1.5%	4.0%	9.0%	8.3%	780
UCM	3.3%	2.4%	12.1%	21.0%	1.3%	2.2%	7.4%	4.5%	7.2%	5.2%	3.8%	5.3%	5.7%	3.8%	3.2%	3.2%	5.0%	5.3%	6.8%	4.9%	5.3%	4.9%	3.1%	718
EHU	4.1%	1.3%	26.9%	6.6%	3.4%	0.7%	17.7%	8.1%	2.9%	11.7%	0.4%	12.4%	2.2%	1.0%	4.7%	5.0%	3.4%	3.1%	14.6%	3.2%	2.5%	5.4%	1.2%	683
UPF	0.7%	4.8%	1.0%	17.6%	2.2%	4.1%	2.6%	5.5%	0.7%	5.8%	3.6%	0.0%	5.8%	4.6%	20.8%	11.3%	5.6%	1.2%	0.7%	1.7%	7.0%	17.6%	0.0%	586
US	5.7%	1.8%	13.8%	10.3%	4.8%	7.1%	18.9%	5.1%	1.8%	5.5%	3.2%	3.4%	10.8%	5.5%	8.7%	1.4%	2.5%	3.9%	3.4%	5.1%	2.5%	6.7%	0.0%	435
USC	5.8%	2.3%	12.6%	20.2%	1.3%	2.8%	7.3%	3.3%	1.5%	4.0%	0.0%	3.5%	5.6%	3.0%	8.1%	3.3%	1.3%	5.3%	21.0%	2.3%	2.5%	3.3%	2.0%	396
HCP-DL (10%)																								
UB	3.4%	2.7%	8.1%	32.2%	0.7%	1.2%	4.3%	5.0%	2.7%	4.5%	3.7%	3.1%	1.1%	2.3%	8.4%	4.0%	8.4%	3.5%	6.8%	4.3%	8.7%	4.7%	2.3%	3030
UAB	1.8%	2.7%	6.7%	29.0%	1.6%	2.2%	6.6%	5.6%	3.7%	4.1%	3.9%	3.6%	2.1%	3.4%	5.1%	3.0%	6.3%	2.8%	12.8%	6.1%	4.4%	6.4%	4.2%	2732
UAM	3.2%	2.4%	11.1%	17.5%	1.5%	1.2%	4.8%	4.0%	0.9%	2.1%	1.8%	4.6%	4.1%	2.0%	7.6%	3.6%	2.9%	2.0%	28.8%	2.3%	3.8%	4.6%	9.5%	1839
UV	7.0%	1.9%	12.5%	19.8%	1.0%	3.2%	5.9%	5.2%	3.7%	3.8%	1.9%	3.8%	2.0%	3.1%	6.4%	2.4%	2.2%	2.3%	25.3%	2.1%	4.7%	6.6%	5.0%	1773
UGR	6.8%	2.3%	7.6%	12.3%	11.9%	2.9%	6.3%	4.6%	6.0%	3.8%	0.7%	2.1%	6.6%	1.2%	4.6%	1.9%	2.6%	2.2%	16.1%	1.9%	5.1%	9.9%	7.4%	1492
UCM	3.4%	2.3%	12.2%	23.4%	2.0%	2.3%	7.6%	3.8%	7.5%	4.3%	4.0%	5.9%	5.2%	4.0%	4.3%	2.8%	5.1%	5.5%	7.9%	4.7%	5.3%	5.0%	3.0%	1409
EHU	3.6%	1.6%	25.6%	8.5%	3.4%	1.6%	15.6%	6.9%	2.6%	8.6%	0.3%	11.5%	3.1%	1.4%	4.5%	4.0%	3.7%	3.2%	15.5%	3.3%	4.4%	6.1%	1.2%	1368
UPF	1.0%	4.5%	1.0%	19.1%	2.6%	4.5%	2.7%	6.0%	0.5%	7.6%	3.7%	0.0%	5.4%	3.6%	21.3%	9.4%	5.1%	1.6%	0.6%	1.8%	6.7%	16.3%	0.0%	1076
US	6.1%	1.5%	14.2%	11.4%	5.9%	5.9%	19.0%	5.7%	1.8%	5.6%	3.1%	4.3%	11.1%	5.2%	8.5%	1.6%	2.3%	3.1%	6.0%	5.2%	2.8%	5.8%	0.0%	879
UPV	8.0%	1.3%	24.6%	3.1%	9.6%	4.5%	22.1%	6.1%	1.5%	2.6%	0.1%	5.1%	9.8%	1.2%	3.1%	3.1%	0.7%	1.1%	8.2%	11.2%	0.7%	4.9%	0.6%	840

Color shades indicates the magnitude (green are universities and subject areas with higher values and red with lower values).

TABLE 4 | Number and percentage of HCP-DL documents by university (ordered by absolute number of HCP-DL documents in the top 1%).

	1%		5%		10%		Rank Abs.			Rank %		
Univ.	Núm. Docs.	%	Núm. Docs.	%	Núm. Docs.	%	1%	var 1-5%	var 5-10%	1%	var 1-5%	var 5-10%
UB	399	1.93%	1594	7.69%	3030	14.62%	1	0	0	4	-2	1
UAB	326	1.73%	1482	7.87%	2732	14.51%	2	0	0	8	4	-3
UV	247	1.88%	961	7.30%	1773	13.46%	3	-1	0	5	-4	0
UAM	207	1.64%	988	7.82%	1839	14.56%	4	1	0	10	5	-1
UGR	173	1.51%	780	6.81%	1492	13.04%	5	0	0	11	1	0
UCM	147	1.01%	718	4.91%	1409	9.63%	6	0	0	21	-2	-7
UPF	137	2.58%	586	11.03%	1076	20.24%	7	-1	0	1	0	0
EHU	124	1.14%	683	6.28%	1368	12.59%	8	1	0	18	5	0
USC	93	1.30%	396	5.54%	825	11.54%	9	-1	-1	13	-5	2
UNIOVI	89	1.45%	394	6.40%	786	12.77%	10	-1	-2	12	1	-1
US	85	0.93%	435	4.78%	879	9.66%	11	2	0	25	-1	-3
ULL	81	1.73%	284	6.08%	549	11.75%	12	-6	0	7	-8	0
UNIZAR	79	0.97%	358	4.41%	792	9.75%	13	-1	2	23	-7	4
UPV	74	0.98%	388	5.12%	840	11.09%	14	2	2	22	0	4
UIB	73	2.13%	283	8.26%	502	14.65%	15	-4	0	3	0	0
UNICAN	72	1.85%	341	8.78%	583	15.00%	16	1	-1	6	4	0
UPC	67	0.89%	368	4.91%	762	10.16%	17	4	-1	30	6	0
URV	67	1.65%	301	7.42%	571	14.07%	18	1	0	9	1	0
UJI	61	2.16%	210	7.43%	414	14.64%	19	-5	-1	2	-5	3
UPM	55	0.73%	321	4.28%	696	9.28%	20	4	1	35	1	-1
USAL	55	1.19%	247	5.33%	438	9.45%	21	1	-3	17	-2	-13
UCO	52	1.25%	218	5.25%	432	10.41%	22	-1	-1	15	-6	-1
UM	41	0.80%	242	4.71%	485	9.44%	23	2	0	34	6	-5
UA	38	0.86%	176	4.00%	368	8.37%	24	-5	0	31	-7	-3
UVIGO	38	0.92%	180	4.37%	385	9.34%	25	-3	2	26	-5	-3
UDG	36	1.19%	186	6.15%	376	12.44%	26	-1	0	16	2	0
UMA	36	0.72%	234	4.65%	490	9.73%	27	5	2	36	7	1
UAH	34	0.97%	187	5.32%	370	10.52%	28	2	-2	24	4	-1
UC3M	32	0.82%	151	3.86%	327	8.36%	29	-1	0	33	-9	0
UNEX	31	1.03%	131	4.33%	236	7.81%	30	-4	-2	19	-13	-14
UCLM	30	0.63%	201	4.19%	467	9.75%	31	6	3	38	3	8
UDL	28	1.30%	138	6.39%	280	12.96%	32	0	-1	14	2	1
UMH	26	0.90%	111	3.84%	255	8.83%	33	-3	2	29	-14	4
URJC	25	0.91%	132	4.82%	297	10.86%	34	1	2	28	3	6
UPO	21	1.01%	124	5.97%	238	11.46%	35	0	0	20	4	-1
UVA	21	0.57%	144	3.92%	296	8.05%	36	5	-1	41	0	-4
UJAEN	18	0.82%	94	4.31%	222	10.17%	37	-3	1	32	-1	10
ULPGC	14	0.65%	89	4.15%	191	8.91%	38	-3	1	37	1	-1
UCA	12	0.48%	102	4.08%	232	9.27%	39	1	0	43	6	1
UDC	10	0.37%	106	3.96%	235	8.78%	40	3	0	46	7	-1
UNAVARRA	10	0.59%	101	5.94%	184	10.82%	41	2	-2	40	23	-3
UNIRIOJA	10	0.92%	43	3.94%	104	9.53%	42	-4	1	27	-13	9
UNILEON	8	0.47%	60	3.51%	138	8.08%	43	0	0	44	-2	2
UAL	7	0.38%	71	3.81%	165	8.85%	44	2	0	45	1	6
UHU	7	0.51%	35	2.53%	104	7.52%	45	-2	3	42	-5	0
UBU	6	0.62%	46	4.75%	97	10.01%	46	2	-3	39	12	2
UPCT	3	0.24%	44	3.52%	102	8.17%	47	2	-1	47	2	2

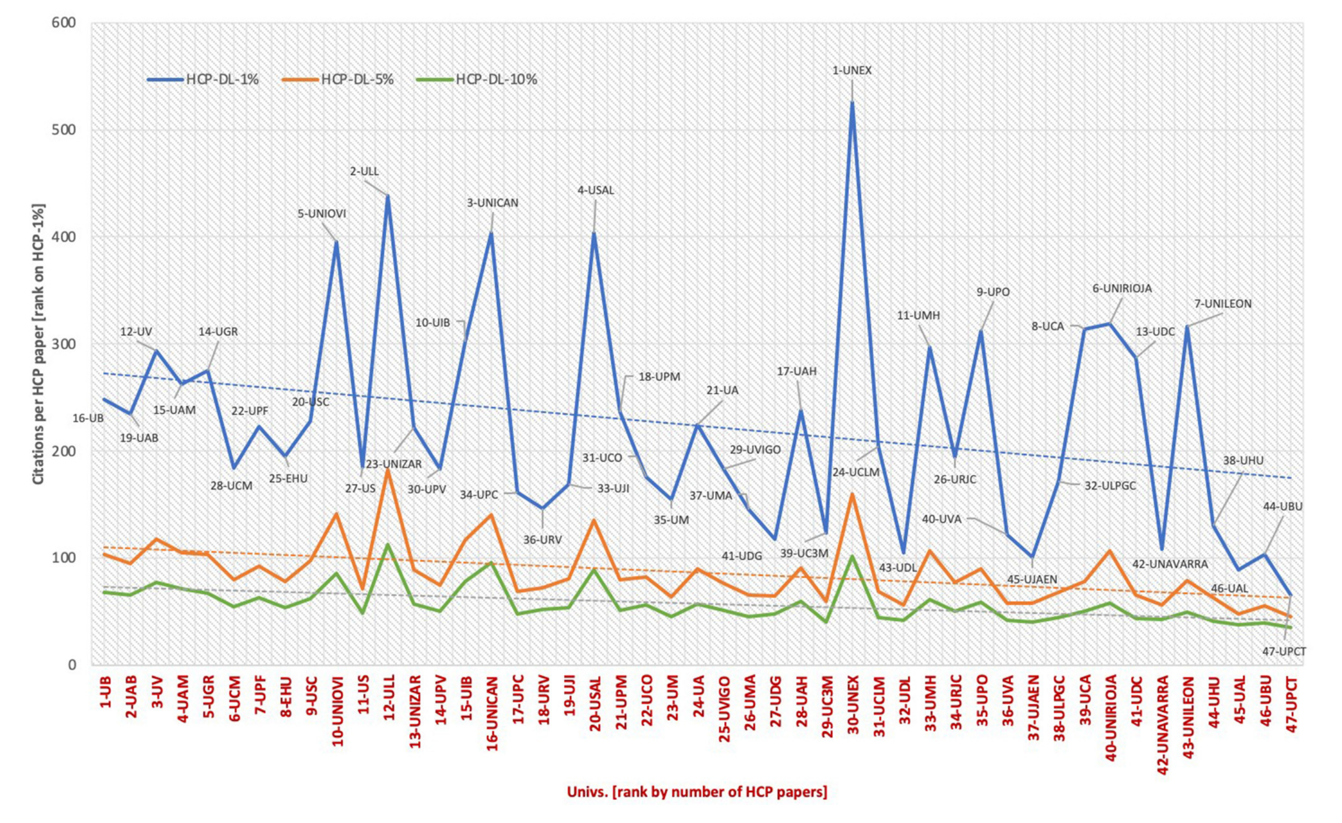

FIGURE 5 | HCP-DL percentage at Spanish public universities (1, 5, and 10%).

if it is published in 2017. The total number of HCP-DLs by area for the 4-year period is shown in APPENDIX III (included as **Supplementary Material**).

The distribution of HCP-DLs by Incites/Essential Science Indicators area in the SUPE universities is presented below, using the limits established in **Table 1**. **Table 2** shows the first 10 universities by absolute number of HCPs in all three top citation groups. This table represents the absolute numbers of HCP-DL for each institution by research area. The values have been colored with a gradient ranging from green (the universities with the highest HCP-DL score) to red (those with the lowest HCP-DL score). The positions change little. The size of universities was determined in base of their number of students (QS Intelligence Unite, 2020) and it can be observed that que the biggest universities (UB, UAB, UV, UAM, UGR, UCM) occupy the leading positions in all three top citation groups. Interestingly, one small university (UPF) places seventh and eighth in the two most-demanding groups. The absolute values for all public universities and Incites/Essential Science Indicators areas are shown in APPENDIX IV (**Supplementary Material**).

The percentage of HCPs by university and area has also been calculated. The values for the first 10 universities by percentages are shown in **Table 3**. The proportions do not always follow the same order as the absolute values. The percentage values by area for all universities are presented in APPENDIX V (**Supplementary Material**).

Once the values for each area are calculated, the total number of HCP-DLs for each university can be found. **Table 4** shows the number of documents that exceed the citation limits at each institution as well as the publishing effort, measured as the percentage of publications that have crossed the citation thresholds out of the total number of documents produced by the university. The last columns show the position (rank) of each institution by its number of documents in the top 1% and by its percentage in the top 1%, together with the changes of position of each institution in terms of rankings and in terms of each citation threshold.

According to the data in **Table 4**, the leading positions are occupied by the large universities, and their order by number of documents remains practically unchanged in the first 10 positions. However, the positions change drastically when publishing efforts (percentage of documents) are compared. With the exception of UPF, which presents the highest HCP-DL percentages in all three top groups, the rest of the universities positioned in the first 10 by number of highly cited documents fall to positions ranging from fourth place for UB to 21st place for UCM.

Figure 5 shows the HCP-DL ratios for public universities at the three citation levels. The universities are presented in order by total number of documents. The positions by publication effort for the top 10% citation group are also presented.

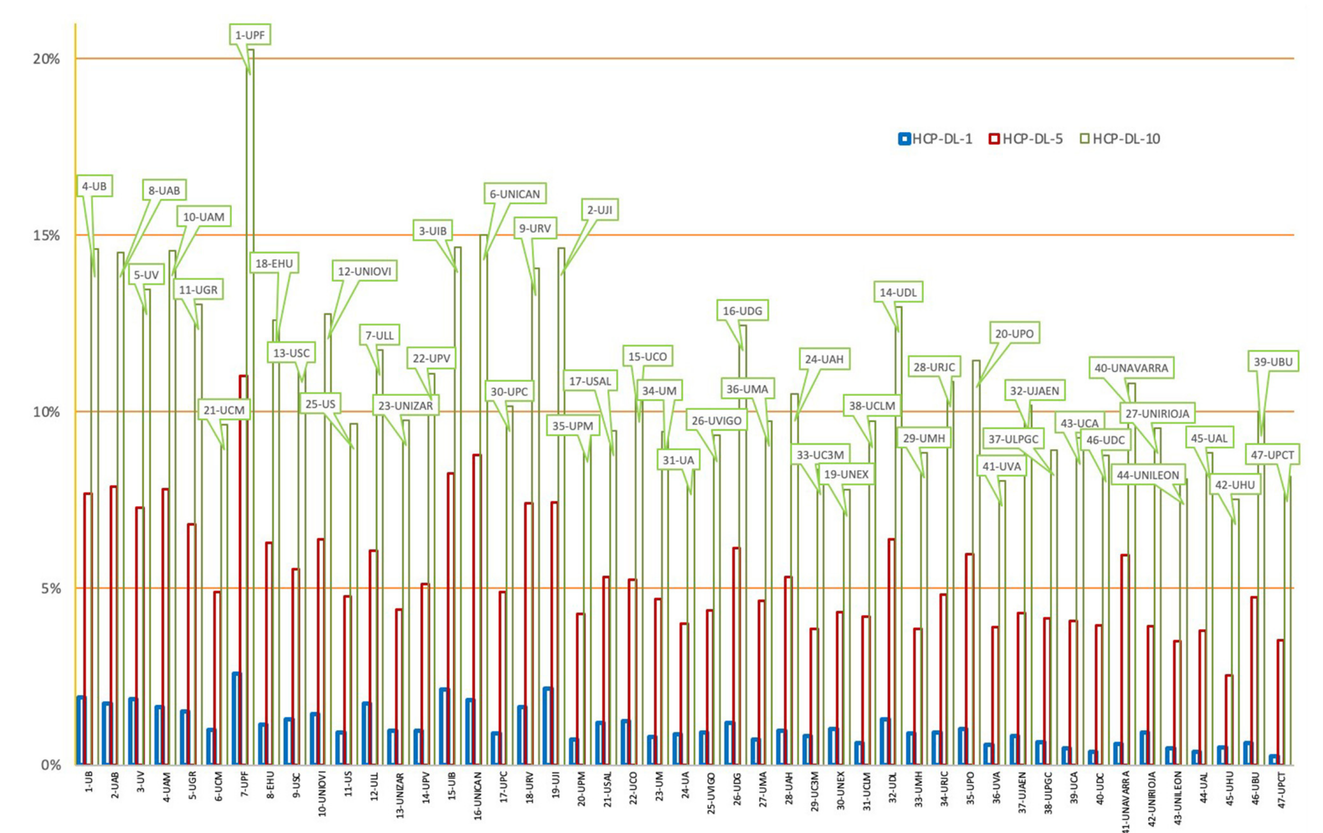

FIGURE 6 | Observed averages of citations per HCP compared to period average thresholds.

Citations per HCP

Once the HCP-DLs have been calculated, a relationship can be found between HCP-DLs and the number of citations received. **Figure 6** shows the comparison between the average number of citations per SUPE document in each Incites/Essential Science Indicators area (Observed averages) and the average per area of the annual thresholds (Threshold average). The data show that the highest averages are in Space Sciences and that Clinical Medicine is in fourth place ahead of Physics.

In addition to revealing the characteristics of the SUPE, these data enable comparisons to be drawn with the information extracted from the WoS HCPs indicator. Although caution is required due to possible differences in the analysis periods or citation windows, some differences can be found. In the case of SUPE, for HCP-DL 1% the thresholds are much higher in Space Sciences (i.e., WoS: 157[1] in 2014, SUPE: 376; or in averages: 120 vs. 238.8) and Multidisciplinary (196 vs. 326), while in Clinical Medicine the average is slightly higher in the case of SUPE than WoS (100.8 vs. 113.5).

By going down to the university level, the number of HCP-DLs can be compared with the citations per HCP document at each institution in the three top citation groups. **Figure 7** shows this relationship by presenting the universities ordered on the abscissa axis by the number of HCP-1% documents (this value appears on the label). To the right of the figure are positioned the universities with the lowest HCP volume, and to the left, those with the highest. Above the trend lines are the universities with higher than expected value in terms of citations per HCP document.

DISCUSSION AND CONCLUSIONS

The present study proposes a methodology focused on identifying HCPs produced by Spanish public universities using the figures for the SUPE as a reference. The proposed indicator, which is termed Highly Cited Paper at the Domestic Level (HCP-DL), provides a new context of comparison that is much better for comparing universities in the same system than the indicators offered by the Web of Science, whose reference is publications world-wide, because the HCP-DL considers the real citation values of the documents published by institutions in the same country.

A number of methodological considerations should be borne in mind. First, to obtain results such as those presented in this paper, there must be a citation window of at least 2 years, to

[1] InCItes Essential Science Indicators. Highly Cited Threshold https://esi.clarivate.com/ThresholdsAction.action.

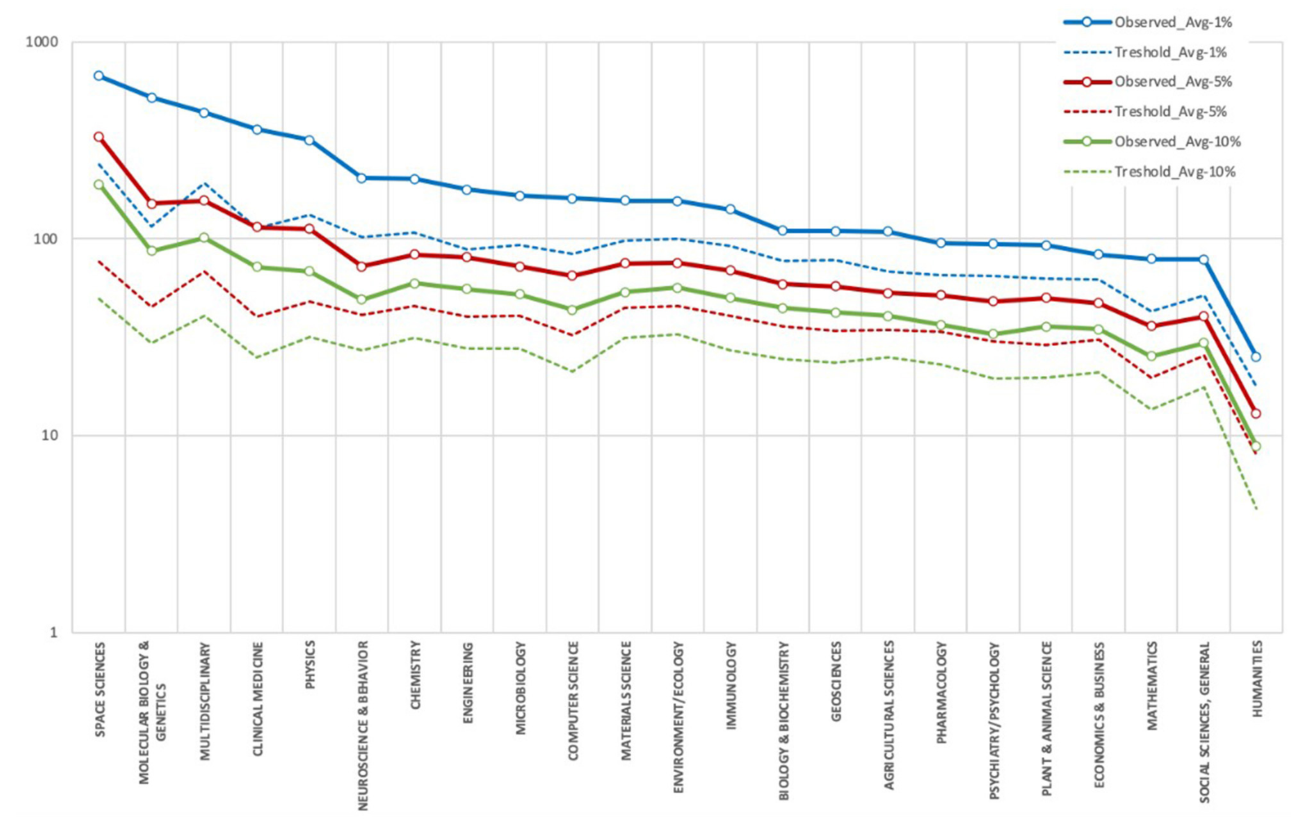

FIGURE 7 | Citation per document by university 2014–2017.

ensure that the citation values of the most-recent publications are consistent with those of the rest of the period. Newly published papers must be allowed time to be cited. That is the reason why the documents analyzed here were published between 2014 and 2017, with citations gathered in February 2020. Future plans are for this methodology to be used to analyse consecutive periods with moving citation windows (2015–2018; 2016–2019; 2017–2020) and thus analyse the development of HCPs by university and by subject area. In addition, the method is easy to use as an additional indicator for evaluating systems like the SUPE due to its relative ease of calculation (as world reference figures are not needed).

Another point is that the comparison makes sense only in the framework of well-defined subject areas. Although Incites/Essential Science Indicators categories have proved adequate, we have included the area of Humanities, which has characteristics of its own; and we believe it is important to differentiate Humanities from Social Sciences, since Humanities has specific citation characteristics that differ from those of many of the social science disciplines. In this way, other authors (Hellqvist, 2010; McManus and Neves, 2020) find that databases such as the Web of Science are too narrow in scope, humanistic scholars publish in their native language and not in English-language journals, and they publish in monographs and anthologies rather than journals he humanities scholars. Another characteristic that also differentiates these researchers from social scientists is that they produce a greater variety of publications, value books, study topics of regional and cultural concerns, and cite much older literature (Huang and Chang, 2008). Therefore, we recommend using this criterion.

The results of the case study of the Spanish university system show a preponderance of HCPs in the field of Space Science. This is because of a tightly clustered small number of Spanish institutions that are members of major international cooperation networks in the category of Astronomy and Astrophysics and publish accordingly. The major non-specialized universities (e.g., UB, UAB, UAM, UGR, UV, UCM) are also observed to have HCPs in many areas, while the polytechnic universities have high visibility in the Computer Science area. It has been observed generally that the presence of HCPs in a given area has to do with a university's teaching and research specialities. For example, UC3M presents domestic HCPs in Economics, Engineering, and Mathematics, but not in areas not covered by its teaching plan, such as the medical sciences.

Obviously, in absolute terms, as the findings of this paper have shown, the big non-specialized universities are major producers of HCPs and hold the leading positions in our results. However, when efficiency is analyzed in relative terms, some small, universities (like UPF, UC3M, UNIOVI, ULL, and UIB) reveal themselves to be more efficient at producing HCPs (%

of HCPs or citations per HCP). Big universities like UB, UAB, UV, UGR, and UAM are also highly efficient and have high HCP percentages in certain categories. Furthermore, there is a large number of universities in the SUPE whose HCP numbers, both absolute and relative, are quite remote from those of the universities mentioned above.

The interest this study has aroused in policymakers, scientific and academic authorities and Spanish accreditation agencies has led us to present the methodology in this special issue on good practices. We believe that, because of its simplicity, its ease of calculation and the knowledge it provides, it can be exported to analyse any country's national systems with a view to ascertaining the impact and visibility of the research done in that country's scientific institutions or in their research subject areas.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

CG-Z: design and implementation of the research, methodological development, data analysis, visualization of data, and writing of the manuscript. SM: computer and methodological development and data extraction. DD: data analysis, writing of the manuscript, and funding acquisition. ES-C: discussion of results, revision of data, and revision of manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was carried out with funds provided by the Rámon y Cajal call (RYC2018-024894-I) from the Spanish Ministry of Science, Research and Universities.

ACKNOWLEDGMENTS

The methodology presented is part of the development of indicators of scientific activity in Spanish universities carried out within the framework of the IUNE Observatory. The IUNE Observatory is supported by the A4U Alliance (Universidad Autónoma de Madrid, Universidad Autónoma de Barcelona, Universidad Pompeu Fabra and Universidad Carlos III de Madrid).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frma.2021.651991/full#supplementary-material

REFERENCES

Aknes, D. (2003). Characteristics of highly cited papers. *Res. Evaluat.* 12, 159–170, doi: 10.3152/147154403781776645

Aksnes, D., and Sivertsen, G. (2004). The effect of highly cited papers on national citation indicators. *Scientometrics* 59, 213–224. doi: 10.1023/B:SCIE.0000018529.58334.eb

ARC (2019). *State of Australian University Research 2018-19*. ERA National Report 2018–19. Available online at: https://dataportal.arc.gov.au/ERA/NationalReport/2018/

Bautista-Puig, N., Mañana-Rodriguez, J., De Filippo, D., Serrano-López, A. E., Lascuráin-Sánchez, M. L., García-Zorita, C., et al. (2020). Actividad investigadora de las universidades españolas (VII). *Informe Annual/Periodo 2009–2018*. Observatorio IUNE. Available online at: http://www.iune.es/

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

Bornmann, L., de Moya Anego', N. F., and Leydesdorff, L. (2012). The new excellence indicator in the world report of the SCImago institutions rankings 2011. *J. Inform.* 6, 333–335. doi: 10.1016/j.joi.2011.11.006

Bornmann, L., Stefaner, M., de Moya Anego', N. F., and Mutz, R. (2014). Ranking and mapping of universities and researchfocused institutions worldwide based on highly-cited papers: a visualization of results from multi-level models. *Online Inform. Rev.* 38, 43–58. doi: 10.1108/OIR-12-2012-0214

Buela-Casal, G., Bermúdez, M. P., Sierra, J. C., Quevedo-Blasco, R., Castro, A., and Guillén-Riquelme, A. (2011). Ranking de 2010 en producción y productividad en investigación de las universidades públicas españolas. *Psicothema* 23, 527–536. Available online at: http://www.psicothema.es/pdf/3918.pdf

Buesa, M., Hieijs, J., and Kahwash, O. (2009). *La Calidad de las Universidades en España. Elaboración de un Índice Multidimensional.* Madrid: Consejo Económico y Social.

Casani, F., De Filippo, D., García-Zorita, C., and Sanz-Casado, E. (2013). Public versus private universities: assessment of research performance, case study of the Spanish university system. *Res. Evaluat.* 23, 48–61 doi: 10.1093/reseval/rvt028

Clarivate Analytics (2020). *InCites Essential Science Indicators. Highly Cited Threshold.* Available online at: https://esi.clarivate.com/ThresholdsAction.action

Corera, E., Chinchilla, Z., De-Moya, F., and Sanz Menéndez, L. (2010). "Producción científica e impacto: ranking general y por áreas de las instituciones universitarias españolas," in *Contribución de las Universidades Españolas al Desarrollo. Informe CYD 2009,* (Barcelona: Fundación CyD), 254–262.

Danell, R. (2011). 'An the quality of scientific work be predicted using information on the author's track record? *J. Am. Soc. Inform. Sci. Technol.* 62, 50–60. doi: 10.1002/asi.21454

De Filippo, D., García-Zorita, C., Lascurain, M. L., Marugán, S., and Sanz-Casado, E. (2014). La actividad investigadora del Sistema Universitario Español (2003-2012). Presentación de los resultados del Observatorio IUNE (III Edición). *Rev. Iberoam. Psicol. Salud* 5, 167–186. Available online at: https://www.redalyc.org/pdf/2451/245131498005.pdf

De Filippo, D., and Sanz-Casado, E. (2018). Bibliometric and altmetric analysis of three social science disciplines. *Front. Res. Metr. Anal.* 3:34. doi: 10.3389/frma.2018.00034

De Filippo, D., Sanz-Casado, E., Casani, F., García-Zorita, C., and Efraín-García, P. (2012). Visibility in international rankings. Strategies for enhancing the competitiveness of Spanish universities. *Scientometrics* 93, 949–966. doi: 10.1007/s11192-012-0749-y

European Commission (2001), *Key Figure 2001. Special Edn. Indicators for Benchmarking of National Research Policies*. Brussels.

Garfield, E. (1955). Citation indexes to science: a new dimension in documentation through association of ideas. *Science* 122, 108–111. doi: 10.1126/science.122.3159.108

Garfield, E. (2006). The history and meaning of the journal impact factor. *JAMA* 295, 90–93. doi: 10.1001/jama.295.1.90

Hellqvist, B. (2010). Referencing in the humanities and its implications for citation analysis. *J. Am. Soc. Inform. Sci. Technol.* 61, 310–318. doi: 10.1002/asi.21256

Hu, Z., Tian, W., Xu, S., and Zhang, C. (2018). Four pitfalls in normalizing citation indicators: an investigation of ESI's selection of highly cited papers. *J. Inform.* 12, 1133–1145 doi: 10.1016/j.joi.2018.09.006

Huang, M. H., and Chang, Y. W. (2008). Characteristics of research output in social sciences and humanities: from a research evaluation perspective. *J. Am. Soc. Inform. Sci. Techno.* 59, 1819–1828. doi: 10.1002/asi.20885

IUNE (2020). *Actividad Investigadora de la Universidad Española. Glosario. Listado de Universidades Públicas y Privadas*. Available online at: http://www.iune.es/es_ES/glosario/listado-universidades.

Lamont, M. (2012). Toward a comparative sociology of valuation and evaluation. *Annu. Rev. Sociol.* 38, 201–221. doi: 10.1146/annurev-soc-070308-120022

Liu, N. C., and Cheng, Y. (2005). Academic ranking of World Universities. *Methodol. Prob. High. Educ. Eur.* 30, 127–136. doi: 10.1080/03797720500260116

Martín-Martín, A., Orduna-Malea, E., and Delgado López-Cózar, E. (2018). Author-level metrics in the new academic profile platforms: the online behavior of the Bibliometrics community. *J. Inform.* 12, 494–509. doi: 10.1016/j.joi.2018.04.001

McManus, C., and Neves, A. A. (2020). Production profiles in Brazilian Science, with special attention to social sciences and humanities. *Scientometrics* 126, 1–23. doi: 10.1007/s11192-020-03452-2

Neylon, C., and Wu, S. (2009). Article-level metrics and the evolution of scientific impact. *PLoS Biol*. 7:e1000242. doi: 10.1371/journal.pbio.1000242

Orduña-Malea, E., Martín-Martín, A., and Delgado-López-Cózar, E. (2016). The next bibliometrics: ALMetrics (Author Level Metrics) and the multiple faces of author impact. *El Prof. Inform*. 25, 485–496. doi: 10.3145/epi.2016.may.18

Priem, J., and Hemminger, M. (2010). Scientometrics 2.0: toward new metrics of scholarly impact on the social web. *First* Monday 15:2874. doi: 10.5210/fm.v15i7.2874

QS Intelligence Unite (2020). *QS Classification. How Do we Classify Institutions?* Available online at: https://www.iu.qs.com/university-rankings/qs-classifications~/

REF (2020). *Ref 2020/02: Guidance on Revisions to REF2021*. Available online at: https://www.ref.ac.uk/media/1417/guidance/on/revisions/to/ref/2021/final.pdf

Ruslan, M. (2003). *The Oxford Handbook of Computational Linguistics*. Oxford: Oxford University Press.

Sanz-Casado, E., De Filippo, D., García-Zorita, C., and Efraín-García, P. (2011). Observatorio IUNE: una nueva herramienta para el seguimiento de la actividad investigadora del sistema universitario español. *Bordón Rev. Orient. Pedagóg.* 63, 101–115. Available online at: https://recyt.fecyt.es/index.php/BORDON/article/view/28977

Sanz-Casado, E., García-Zorita, C., Serrano-López, A., Efraín-García, P., and De Filippo, D. (2013). Rankings nacionales elaborados a partir de múltiples indicadores frente a los de índices sintéticos. *Rev. Españ. Document. Cient.* 36:e012. doi: 10.3989/redc.2013.3.1.023

Seglen, P. O. (1997). Why the impact factor of journals should not be used for evaluating research. *Brit. Med. J.* 314, 498–502. doi: 10.1136/bmj.314.7079.497

Sivertsen, G. (2018). The Norwegian model in Norway. *J. Data Infor. Sci*. 3, 3–19. doi: 10.2478/jdis-2018-0017

Szomszor, M., Adams, J., Fry, R., Gebert, C., Pendlebury, D. A., Potter, R. W. K., et al. (2021). Interpreting bibliometric data. *Front. Res. Metr. Anal*. 5:628703. doi: 10.3389/frma.2020.628703

Tijssen, R., Visser, M., and van Leeuwen, T. (2002). Benchmarking international scientific excellence: are highly cited research papers an appropriate frame of reference? *Scientometrics* 54, 381–397. doi: 10.1023/A:1016082432660

Torres-Salinas, D., Delgado-López-Cózar, E., García-Moreno-Torres, J., and Herrera, F (2011). Ranking ISI de las universidades españolas según campos científicos: descripción y resultados. *El Prof. Inform.* 20, 111–118. doi: 10.3145/epi.2011.ene.14

van-Raan, A. (2000), "The Pandora's box of citation analysis: measuring scientific excellence — the last evil?," in *The Web of Knowledge: A Festschrift in Honor of Eugene Garfield,* eds B Cronin and H B Atkins (Medford: ASIS), 301–319.

van-Raan, A. (2005). Fatal attraction: conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics* 62, 133–143. doi: 10.1007/s11192-005-0008-6

Waltman, L., Caleron6etricsction: conceptua Noyons, E. C. M., Tijssen, R. J. W. van Eck, N. J., et al. (2012). The Leiden Ranking 2011/2012: data collection, indicators, and interpretation. *J. Am. Soc. Inform. Sci. Technol.* 63, 2419–2432. doi: 10.1002/asi.22708

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Research Metrics and Analytics

ORIGINAL RESEARCH
published: 30 June 2021
doi: 10.3389/frma.2021.696470

Check for updates

Bibliometric Reports for Institutions: Best Practices in a Responsible Metrics Scenario

Alvaro Cabezas-Clavijo[1] and Daniel Torres-Salinas[2]*

[1]Faculty of Business And Communication Studies, Universidad Internacional de La Rioja (UNIR), Logrono, Spain, [2]Department of Information And Communication Sciences, Universidad de Granada, Granada, Spain

Carrying out bibliometric reports is one of the common tasks performed by librarians and practitioners within the framework of their professional duties. The emergence of novel data sources, the need to measure new research activities and the growing demand for fairer and more equitable evaluation within the framework of the Responsible Metrics movement has led to calls for a review of the traditional approaches to these types of reports. The main goal of this study is to outline a series of recommendations for bibliometricians, consultants and research support librarians when drafting bibliometric reports in their institutions. These best practices can significantly enhance the quality and utility of bibliometric reports, posing their practitioners as key players in the science management process.

Keywords: bibliometrics, reports, best practices, Responsible metrics, Responsible research and innovation

Edited by:
Juan Ignacio Gorraiz, Universitat Wien, Austria

Reviewed by:
Ursula Ulrych, University of Vienna, Austria
Nicola De Bellis, University of Modena and Reggio Emilia, Italy

****Correspondence:***
Alvaro Cabezas-Clavijo
alvaro.cabezas@unir.net

Specialty section:
This article was submitted to Research Assessment, a section of the journal Frontiers in Research Metrics and Analytics

Received: *16 April 2021*
Accepted: *14 June 2021*
Published: *30 June 2021*

Citation:
Cabezas-Clavijo A and Torres-Salinas D (2021) Bibliometric Reports for Institutions: Best Practices in a Responsible Metrics Scenario. Front. Res. Metr. Anal. 6:696470. doi: 10.3389/frma.2021.696470

INTRODUCTION

In recent years, the evaluation of the performance of research institutions has become an increasingly complex task for universities, research centers and funding and evaluating bodies around the world. The emergence of novel data sources, the measurement of new research activities beyond the mere publication of scientific results and the increasing need for fairer, more equitable and responsible assessment procedures have led to a new scenario characterized by multidimensional evaluations that consider aspects such as knowledge transfer, the diversity of research outputs that an institution can generate and other ethical, integrity and equity issues. These aspects call for a rethinking of the traditional bibliometric reports, i.e., those that mainly analyze results in scientific journals and use citation indexes such as Web of Science or Scopus, which are produced or commissioned by research institutions (for example, Universidad de Granada, 2014; Barcelona Institute of Science and Technology, 2019).

Bibliometric Units

The growing demand for proven bibliometric information and the increasing complexity of research measurement processes has generated the appearance in R and D centers and universities of departments specializing in the evaluation of scientific activity, the so-called 'bibliometrics units' or 'science evaluation units', among other names. These units may be configured in very different ways, with very different roles and tasks depending on the needs of each institution. The functions performed by these units include (Torres-Salinas and Jiménez-Contreras, 2012): a) management of research information sources b) generation of analysis, prospective and surveillance reports and c) training, advice and expert consultation. **Table 1** highlights some of the bibliometrics units that have been created in recent years in Spain, following in the footsteps of the pioneering Bibliometrics Department of the University of Vienna launched in 2009.

TABLE 1 | Examples of Bibliometrics Units in Spanish universities. Source: Own Elaboration.

University	Name of the department/unit	Year
University of Granada	Unidad de Evaluación de la Actividad Cientifica (Scientific Activity Evaluation Unit)	2011
University of Las Palmas	Unidad de Bibliometria (Bibliometrics Unit)	2013
University of Navarre	Unidad de Bibliometria (Bibliometrics Unit)	2014
University of Seville	Unidad de Bibliometria (Bibliometrics Unit)	2018
University of Cadiz	Unidad de Bibliometria (Bibliometrics Unit)	Not Av.
University of the Basque Country	Unidad de Bibliometría–Observatorio de la Produccion Científica (Bibliometrics Unit - Scientific Production Observatory)	Not Av.

One of the most important tasks of practitioners (research support librarians, research analytics librarians, liaison librarians, research performance analysts, bibliometrics officers, consultants, bibliometricians, etc.), whether in the framework of higher education institutions or working in consulting firms, is the preparation of bibliometric reports. These quantitative reports tend to have a descriptive purpose, that is, they aim to reflect the state of the research at a given moment, for example in a university, or an evaluative purpose, for example if the report is used to support the assessment of a certain funding call or area of the institution.

The Department for Bibliometrics and Publication Strategies of the University of Vienna is a good example of a unit which prepares both descriptive and evaluative reports, using its own methodology (Gumpenberger et al., 2012; Gorraiz et al., 2020). Similar activities are carried out by different institutions across the world, such as the University of New South Wales in Australia (Drummond and Wartho, 2009), the Technical University of Munich in Germany (Leiß 2017) and Universidad San Ignacio de Loyola in Peru (Pacheco-Mendoza et al., 2020). A special case is the Center for Science and Technology Studies (CWTS) at Leiden University in the Netherlands, which has become a key provider of bibliometric assessment reports for a wide range of institutions at an international level (Petersohn and Heinze, 2018) through its company CWTS BV. It is also necessary to highlight the role that numerous consulting firms have played in the preparation of bibliometric and evaluation reports, such as Science-Metrix, Technopolis, Evidence LTD, Digital Science Consultancy, EC3metrics and the Institute for Scientific Information (ISI), re-established in 2018 as the analytics expertise service of Clarivate Analytics.

Whether consulting firms or bibliometrics units, the preparation of these kinds of documents requires a number of specific skills (Iribarren-Maestro, 2018): knowledge of the different publication and citation guidelines in the different scientific areas; application of knowledge regarding statistics, scientific policies, legislation and other matters to the analysis and interpretation of results; recognition of the characteristics of the publications of scientific journals and publication models; identification of the characteristics of editorial quality products, and insights into the different university rankings depending on the nature of the reports requested. According to the competency model for those supporting bibliometrics (Cox et al., 2019), tasks associated with the design and execution of bibliometric reports–such as evaluation of departmental/research center performance, or assessment of institutional performance–are considered as 'specialist tasks' by the professional community, the highest level of competency for bibliometric work (Cox et al., 2019).

Depending on the needs of each institution, different types of reports may be generated (Torres-Salinas and Jiménez-Contreras, 2012; Gorraiz et al., 2020):

- Bibliometric reports at an institutional level: the results of these reports may be included in annual reports, with the main goal being to provide a precise overview of the state of the research at a particular point in time.
- Case studies: bibliometric reports focusing on a certain aspect of the research which is of strategic interest to the institution. For example, they may focus on a specific topic (gender, collaboration, open access) or a specific area (engineering, arts, fine arts, biomedicine).
- Decision-making and supporting reports: these provide useful information for scientific policymakers, such as reports for specific funding calls, faculty evaluations, recruitment or appointment procedures.
- Informative bibliometric reports: intended for the dissemination of key research findings by the University Communications Office to the media and the general public.

The Responsible Metrics Scenario

Bibliometric professionals should also be aware of the Responsible Metrics movement and associated international manifestos and recommendations calling for the responsible use of bibliometric indicators. This matter has been gaining repercussion in recent years and more and more institutions are integrating some of these fundamental principles in their evaluation policies. The two main documents defining the responsible use of evaluation indicators are the San Francisco Declaration on Research Assessment (DORA, 2012) sponsored by the American Society for Cell Biology and the Leiden Manifesto for Research Metrics (Hicks et al., 2015) issued by a number of renowned bibliometric experts. These documents call for a more balanced and fairer approach to the use of metrics in research evaluation, especially with regard to individuals (recruitment, staff promotion, scholarships, calls for mobility, grants, etc.). Bibliometric practitioners cannot ignore this perspective when designing and executing these types of studies, even though it could be

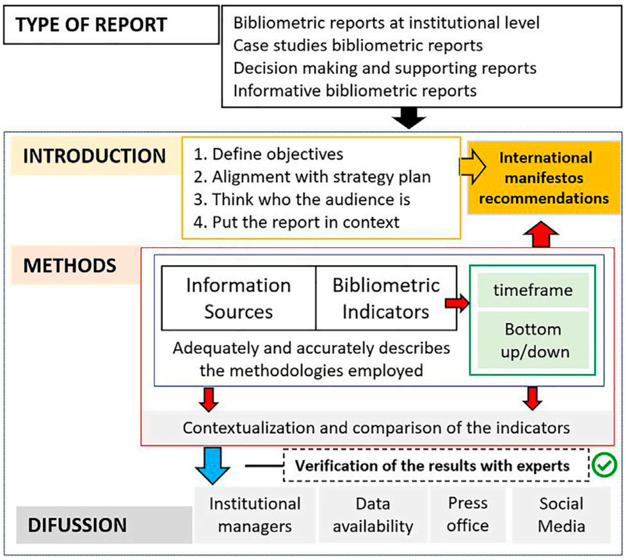

FIGURE 1 | "Flowchart of the main processes and decisions for bibliometric reporting at an institutional level in a Responsible Metrics scenario".

argued it is "time-consuming, expensive and requires a significant increase in bibliometric expertise" (Coombs and Peters, 2017). It is especially relevant for practitioners affiliated with institutions which are signatories of these manifestos. According to the 2019 Responsible Metrics State of the Art Survey (Robinson-García and Gadd, 2019), 23% of the respondents belonged to institutions that have signed DORA.

Purpose of This Study

In light of this professional scenario, the main purpose of this study is to establish a series of best practices and recommendations for bibliometricians, consultants and research support librarians when drafting bibliometric reports for their institution. These guidelines are intended mainly for generation of bibliometric reports at an institutional level and case studies, although some of them are also applicable to evaluative reports to support decision-making. These recommendations are based on an extensive number of reports generated by different universities, consulting firms and bibliometric experts and are also guided by the framework offered by the Responsible Metrics principles. **Figure 1** provides a concise summary of the reporting process and the decisions we will need to make in order to prepare bibliometric reports. In ***Best Practices for Bibliometric Reports*** we explain each of these actions and processes in greater detail.

BEST PRACTICES FOR BIBLIOMETRIC REPORTS

A series of recommendations have been set out below as a guideline to follow when preparing a bibliometric report. They include different international recommendations and should be considered as a compendium of best practices with a special emphasis on bibliometric reports for R and D institutions. These ten recommendations may be divided into three different blocks. The first block includes preliminary aspects that introduce the report to the reader and is concerned with adequate definition of the objectives and correct introduction of the socioeconomic context of our institution. The second block compiles four recommendations relating to all the methodological aspects of the report. We will have to make multiple decisions, starting with the databases and indicators we are going to use. This block also includes advice on the importance of adequately describing the methods and contextualization/comparison of the results. Finally, the third block presents four best practices that are based on Responsible Metrics principles and the growing demand for transparency and accountability in modern society.

Preliminary Matters

Define the Objectives

Any report must indicate the objectives of the analysis carried out, contextualizing it within the framework of other similar studies carried out by the same institution. It should also be indicated whether the report is regular (biannual, annual) or if it is part of a series. The orientation of the report (descriptive or evaluative) and the purpose of the study must be adequately broken down. It is essential for it to be duly aligned with the objectives of the institution, with the purpose of the report being linked to the goals designed in the strategic plan of the organization.

For example, if one of the objectives of the institution is to expand its international presence, this purpose may be matched to indicators referring to international publications or collaborations. The need for the use of metrics should be adequately explained, since it should not be overlooked that in certain contexts the use of bibliometric indicators may be seen "as a challenge to academic freedom and to the university's traditional role as a center in society of critical and independent thinking" (Cox et al., 2019). Likewise, the target audience of the study should be indicated (research managers, media, wider public, institution staff), along with the use that may be made of it and the context in which the information included in the report may be used.

Provide a Socioeconomic Context for the Institution

Offer a context for the results presented. It is a good idea to devote a brief introductory chapter to the socioeconomic aspects of your organization to facilitate an understanding of the bibliometric indicators used. For example, information could be included on GDP, labor structure, employment rates, production sectors, R and D investment, university staff, students, etc. This context may explain or at least qualify and generate a better understanding of the results obtained. This contextual information is especially important for readers unfamiliar with the institution or who do not belong to its sphere of influence. For example, in the case of reports on university alliances, international research networks or multicenter research, a brief description of the social and economic environment of each institution can provide valuable information about the achievements reached by their

TABLE 2 | Example of how indicators can be defined and described in a bibliometric report. Source: Own Elaboration based on Karolinska Institutet (2014).

Designation	Hirsch index
Abbreviation	H-Index
Definition	The h-index is the number of publications (h) attributed to the unit analyzed during the time span analyzed that have at least h citations.
Calculation and/or Formula	Find the unit's published articles in a citation index and sort them in descending order by number of citations. Count articles from the top of the list downwards and when the number of an article rises above the citation count for that same article, the number of the preceding article is to be counted as the h-index.
Data Requirements	Requires data from a comprehensive citation database (Web of Science, Scopus or Google Scholar)
Advantages	➔ Very easy to calculate in different databases ➔ Included in different research profiles (Google Scholar, Scopus ID, …) ➔ Accepted and very well known by the scientific community
Disadvantages	➔ h-index gives positive bias to senior researchers with older articles ➔ The indicator is not field-normalized, which makes it unsuitable for. ➔ comparisons between researchers in different research fields
Use and application	We use the h-Index to generate author rankings and detect the researchers with the greatest impact in different areas.
Reference	Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences of the United States of America, 102(46), 16569–16572.

components, since the starting points and goals of each node within the network may be very different.

Methodological Aspects

Select and Describe the Used Indicators

One of the key aspects of a report is to determine which indicators are best suited to achieve the objectives. Define a set of indicators that measure different dimensions; reports that assess just one dimension of scientific activity, such as publications in top-ranked journals, without considering other variables (scientific impact, collaboration, training capacity, research funds, etc.) should be avoided. There are hundreds of indicators that allow us to offer a multidimensional view of the research. However, only use indicators which are validated by the scientific community through publication in peer reviewed outputs, and which are broadly used by bibliometric experts. . Use also metrics which are easy to interpret, as non-experts have difficulty understanding complex indicators. Avoid inventing your own indicators, especially composite metrics that mix several indicators in a single measure. Likewise, avoid conscious attempts to manipulate the results, for example choosing metrics that may clearly favor your institution or certain areas or researchers within it.

Always include in the institutional report a precise definition of any of the indicators you are using, describing particularly detailed calculations and/or formula and their advantages and shortcomings. **Table 2** offers an example of how to describe the indicators. You can also draw inspiration from handbooks such as "The Evaluation of research by Scientometric Indicators" (Vinkler, 2010), "Applied Evaluative Infometrics" (Moed, 2017) or "Handbook of Bibliometric Indicators: Quantitative Tools for Studying and Evaluating Research" (Todeschini and Baccini, 2017) to help you choose the right indicator. Karolinska Institutet offers a good example of best practices for the description and use of bibliometric indicators at an institutional level. It would be desirable for all institutions to have documents like the "Bibliometric Handbook for Karolinska Institutet" (Rehn and Kronman, 2008) and "Bibliometric Indicators—Definitions and Usage at Karolinska Institutet" (Karolinska Institutet, 2014).

Fortunately, nowadays most of the indicators we need can be found and calculated in the most popular bibliometric databases. Commercial suppliers (Clarivate Analytics, Elsevier, etc.) propose a large number of indicators in SciVal[1] and InCites[2] handbooks. In both cases, definitions, calculations and formulas are presented. The metrics offered on these platforms highlight the huge number of bibliometric indicators available. InCites has a total of 64 indicators classified into six sections (Productivity, Impact, Collaboration, Reputation, Open Access and Author Position). On the other hand, SciVal offers 29 bibliometric indicators classified into seven groups (Collaboration, Published, Viewed, Cited, Economic Impact, Societal Impact, and Awards). In the case of SciVal, mention should also be made of the Snowball Metrics Initiative (Colledge, 2017), which develops a set of standard methodologies to calculate research metrics in a consistent way regardless of the data sources.

Bibliometricians can also take advantage of new indicators such as altmetrics and social media metrics offered by a number of platforms, as they can provide valuable information to study new forms of interaction between the general public, scholars and academic stakeholders (Zahedi and Costas, 2018) and measure the broader impact of research. Bornmann (2014) identifies four benefits of altmetrics compared to traditional metrics: broadness, diversity, speed, and openness. Nevertheless, serious concerns have arisen regarding the meaning of these metrics and a number of limitations may also be identified concerning the data quality, such as bias, measurement standards, normalization and, replication (Bornmann, 2014).

Use the Appropiate Sources, Databases and Tools

Use a diverse range of databases, avoiding the use of single sources that show significant results only for a limited number

[1]https://service.elsevier.com/app/answers/detail/a_id/13936/supporthub/scival/
[2]http://help.prod-incites.com/inCites2Live/8980-TRS/version/default/part/AttachmentData/data/InCites-Indicators-Handbook-6%2019.pdf

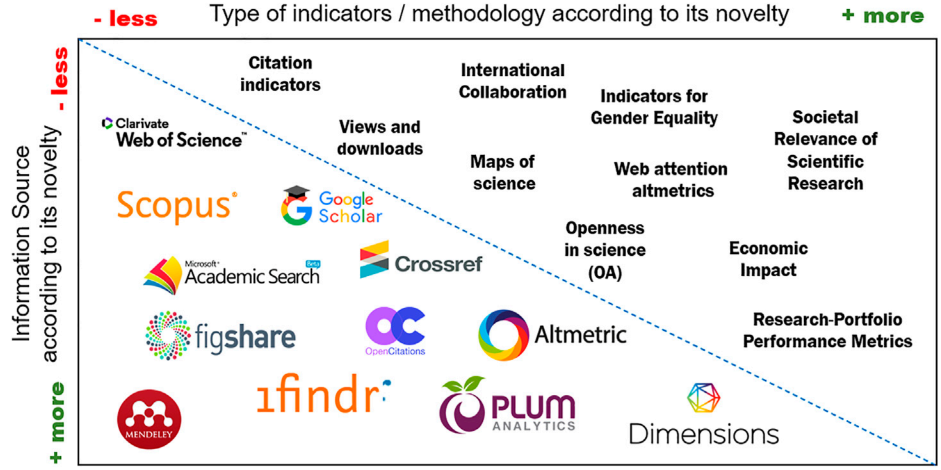

FIGURE 2 | "Main information sources and indicators for bibliometric reports".

of disciplines. Local or national bibliometric products should be used to complement areas that are not well covered by international databases, as occurs in the arts and social sciences. According to Jappe (2020), only one out of every four bibliometric assessment studies uses national sources. Current Research Information Systems (CRIS), as well as institutional administrative databases and other non-bibliometric sources, can offer a more precise picture of research in the institution and are critical to offer accurate and significant results. Nevertheless, using CRIS and internal databases (in relation to grants or human resources, for example) may require intense work with the institution's administrators and a time-consuming curation process.

News databases and altmetric sources (e.g., PlumX[3] and Altmetric.com[4]) can provide relevant information on the outreach and communication activities of the institution and its social/societal impact, while university rankings (e.g. Leiden Ranking[5], ARWU[6] and Webometrics[7]) can provide information on the institution's research impact and web visibility. The report work team should also be aware of the possibility of automatically collecting data from various sources *via* API. **Figure 2** offers an overview of some of the information sources currently available and the indicators they allow us to calculate.

Another central issue to be determined by practitioners is the software used for data gathering and presentation of results. There are a number of bibliometric suites on the market developed by renowned companies such as SciVal[8] (Elsevier), InCites[9] (Clarivate) and Dimensions Analytics[10] (Digital Science) that allow generation of results in various formats. There are also several free products that can also be helpful when preparing the full report or completing a specific section of it, such as Publish or Perish[11], Bibliometrix[12], Vosviewer[13], and Scimat[14], so these should be thoroughly evaluated. Moral-Muñoz, et al. (2020) provides a valuable review of the various tools available for conducting bibliometric and scientometric analyses.

Control the Methods

Clearly define the methodological aspects: chronological framework, approach, units of analysis, data collection methods, databases used, coverage, etc. The reports published by CWTS (2017) and the Nordic Institute for Studies in Innovation, Research and Education (NIFU, 2019) clearly address these issues. Point out the limitations of the study so that the results may be properly contextualized. Remember that it should be possible to reproduce and replicate all the aspects of the study. For example, regarding the unit of analysis, the aggregation level used should be stated. Three levels may be distinguished: a) micro-level, when the report focuses on individual researchers or research groups, b) meso-level, when it refers to departments or institutions, and c) macro-level, when the assessment is related to a region or country.

A further consideration is the way the indicators are compiled, i.e., bottom-up or top-down. Under the bottom-up approach, analysis begins with the data collection of the individual researchers of the institution (micro level) before moving up to higher aggregation levels by grouping the documents. This technique requires great precision in the compilation as well as verification by the researchers evaluated, and is the recommended procedure in reports

[3]https://plumanalytics.com/learn/about-metrics/
[4]https://www.altmetric.com/
[5]https://www.leidenranking.com/
[6]http://www.shanghairanking.com/
[7]https://www.webometrics.info/
[8]https://www.scival.com/
[9]https://incites.clarivate.com/
[10]https://www.dimensions.ai/products/dimensions-analytics/
[11]https://harzing.com/resources/publish-or-perish
[12]https://www.bibliometrix.org/
[13]https://www.vosviewer.com/
[14]https://sci2s.ugr.es/scimat/

aimed at decision-making by research managers. This approach also allows retrieval of documents that researchers have produced outside their current work centers (Costas, 2008). On the other hand, under the top-down approach the data is collected at an institutional level and the analysis may then be lowered to other more disaggregated levels. Data collection under this approach is much faster (since a search by affiliation can be done in selected data sources), although it lacks the accuracy and precision of the former approach, making it more appropriate for descriptive studies.

Use relatively long timeframes to observe the evolution of the indicators over time. A minimum period of five years of analysis is recommended. The use of short timeframes (two to three years) could overestimate some indicators which may be affected by a specific event or by specific legislation or regulations, thereby not duly reflecting the evolution and dynamics of a particular aspect of research within the institution. A useful technique to improve the stability of indicators that avoids changes in trends caused by a specific event is the use of overlapping periods (for example, 2017–2019; 2018–2020; 2019–2021). Likewise, caution must be exercised with the data of the most recent year, since they may be affected by updating procedures in the data sources, as well as by insufficient volume of information (e.g., citation window). Finally, we recommended maintaining stability over time in the methodologies used. In the case of annual reports, the same set of basic indicators should be used and avoid changing the data providers in order to facilitate comparability of annual trends.

Compare and Contextualize the Results

Always compare the results obtained with other institutions and contextualize them by region, country or thematic area in order to determine and understand the performance of your center. The use of comparisons and contextualization is of key importance to take full advantage of bibliometric information. Comparisons should be made with institutions with similar profiles, i.e., analogous size, objectives and disciplines. For example, a historical university with a general profile should not be compared with a technical university or a recently established center focusing on biomedical sciences. Use international benchmarks to contextualize the performance of the university or center such as Essential Science Indicators, or statistics reported by organizations such as the Organization for Economic Cooperation and Development (OECD), the United States National Science Foundation (NSF) or Eurostat at a European level. General and disciplinary baselines can be used to assist with in-depth interpretation of the information. **Table 3** shows a real example of benchmarking for the University of Granada with three indicators, one absolute and two relative indicators.

Responsible Metrics Issues

Obtain Validation

Early drafts should be revised by a scientific committee of experts working in your institution, which can provide useful insights to improve the quality of the report and detect possible errors and inconsistencies. You can also ask for the support of policymakers or relevant researchers from different disciplines who can explain and qualify specific results involving unique publication and citation habits, or anomalous data, which may be determined by aspects relating to the sources used, legislative changes or socioeconomic conditions. When dealing with sensitive topics or especially relevant issues, an expert committee can be set up to guide and validate the data, methods, and procedures.

Pay Attention to Diversity

Be aware of the diversity of research areas present in the institution; avoid solely applying indicators intended for experimental or biomedical sciences. Consider research in local languages as well as activities that contribute to improve the socioeconomic environment in the area around the University or center analyzed. Avoid solely paper-focused reports. Bear in mind the Hong Kong Principles for Assessing Researchers (Moher et al., 2020) and try to introduce indicators aimed at valuing a broader range of research and scholarship, such as replication, innovation, translation, synthesis, and meta-research, peer review, mentoring, outreach, and knowledge exchange, among others.

Apply Ethical, Integrity and Equality Principles

Apply ethical, integrity and equality principles in accordance with the numerous international recommendations in this regard. Consider the latest developments in Responsible Research and Innovation and try to incorporate some of these new indicators in your analysis. For example, the SUPER MoRRI (Scientific Understanding and Provision of an Enhanced and Robust Monitoring system for Responsible Research and Innovation) Project[15] identifies up to 36 indicators in six different areas: gender equality, literacy and science education, public engagement, ethics, open access and governance. Finally, any conflicts of interest that may arise should be disclosed.

Make the Report Public and Open Your Data

Make the results of the report available to the public, unless there is a confidentiality agreement to restrict the publication. Present the data in an attractive way through interactive reports, infographics or dedicated websites. For example, the LiveMetrics project[16] of the University of Granada presents bibliometric indicators and R and D statistics for the University in a dynamic and up-to-date way (**Figure 3**). Also take advantage of general and academic social media and the University Communications Office to maximize the reach of your report. Make the raw data of the reports open and accessible to facilitate the replicability of the study and its reuse by other researchers. Upload your

[15]https://www.super-morri.eu/super-morri/index.php

[16]https://livemetrics.ugr.es/

TABLE 3 | Example of contextualization and comparison of bibliometric indicators at an institutional level: University of Granada. Source: Own Elaboration.

1.A. Example of contextualization of the University of Granada with three baselines	Web of Science Documents	Category Normalized Citation Impact	% Documents in Q1 Journals
University of Granada	21,312	1.26	53.11%
Global Baseline	15,834,230	0.96	47.33%
EU-27 Baseline	4,031,472	1.1	50.20%
Spain—Baseline	548,508	1.2	55.46%
1.B. Example of comparison of the University of Granada with three universities	**Web of Science Documents**	**Category Normalized Citation Impact**	**% Documents in Q1 Journals**
University of Barcelona	45,919	1.68	62.70%
University of Granada	21,312	1.26	53.11%
University of Seville	18,890	1.05	53.40%
Complutense University of Madrid	26,902	1.12	52.29%

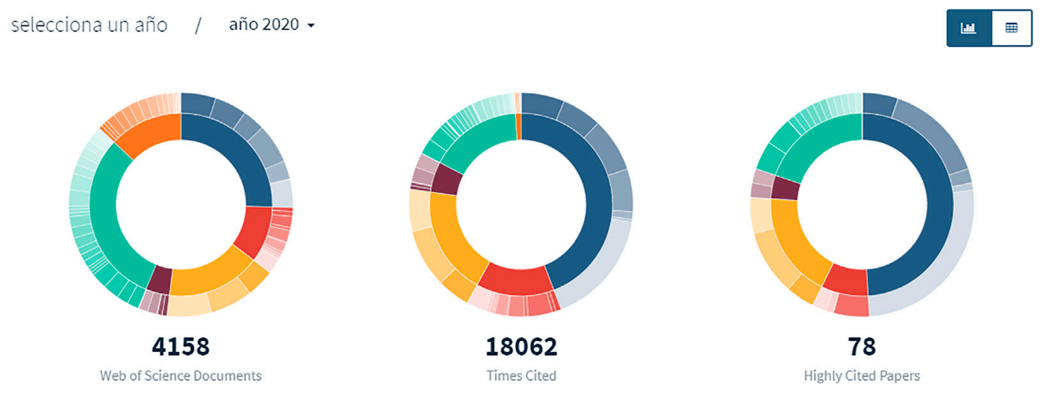

FIGURE 3 | "Example of data visualization–LiveMetrics project".

data to the Open Data platform of your institution or use an external repository.

FINAL REMARKS

The preparation of more responsible bibliometric reports within the framework of scientific policies that seek to be increasingly fair and equitable and more closely connected with the challenges of modern society constitutes a major challenge for librarians and evaluation specialists. This study has presented a series of recommendations for a new generation of bibliometric studies that definitively abandon dependence on single sources and the exclusive measurement of scientific articles, in favor of a broader vision that adequately evaluates the different forms of research carried out by universities and R and D centers.

We are aware that very few reports will be able to take into account all the variables suggested in this study, nonetheless the possibility exists for these types of analyses to move forward in the direction set by new trends in the responsible metrics scenario. The more professionals assume and implement these best practices, the greater the influence they will have in the science management process, offering relevant answers to the challenges posed by research activity today.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

This work has been possible thanks to financial support from "InfluScience - Scientists with social influence: a model to measure knowledge transfer in the digital society" (PID 2019-

109127RB-I00/SRA/10.13039/501100011033), a project funded by the scientific research team grants from the Ministry of Science and Innovation of Spain. Daniel Torres-Salinas has received funding from the University of Granada's "Plan Propio de Investigación y Transferencia" under the "Reincorporación de Jóvenes Doctores" grant.

REFERENCES

Barcelona Institute of Science and Technology (2019). *Bibliometric Report of the Scientific Production Barcelona Institute of Science and Technology (BIST)*. Available at: https://bist.eu/wp-content/uploads/2019/07/4b.-Informe-de-Bibliometria.pdf (Accessed June 11, 2021)

Bornmann, L. (2014). Do altmetrics point to the Broader Impact of Research? an Overview of Benefits and Disadvantages of Altmetrics. *J. Informetrics* 8, 895–903. doi:10.1016/j.joi.2014.09.005

Colledge, L. (2017). *Snowball Metrics Recipe Book*. Available at: https://www.elsevier.com/__data/assets/pdf_file/0006/53169/Snowball_Metrics_Recipe_Book.pdf (Accessed April 5, 2021)

Coombs, S. K., and Peters, I. (2017). The Leiden Manifesto under Review: what Libraries Can Learn from it. *Dlp* 33, 324–338. doi:10.1108/DLP-01-2017-0004

Costas, R. (2008). *Análisis bibliométrico de la actividad científica de los investigadores del CSIC en tres áreas, Biología y Biomedicina, Ciencia de los Materiales y Recursos Naturales: una aproximación metodológica a nivel micro (Web of Science, 1994-2004)* [Doctoral dissertation]. Madrid: Universidad Carlos III de Madrid).

Cox, A., Gadd, E., Petersohn, S., and Sbaffi, L. (2019). Competencies for Bibliometrics. *J. Librarianship Inf. Sci.* 51, 746–762. doi:10.1177/0961000617728111

CWTS (2017). *Research performance analysis for the SOR programme of the Rijksinstituut voor Volksgezondheid en Milieu (RIVM) 2011-2014/15*. Available at: https://www.rivm.nl/sites/default/files/2018-11/CWTSbibliometricreport%20SOR.pdf (Accessed April 5, 2021)

DORA (2012). *Declaration on Research Assessment (DORA)*. Available at: https://sfdora.org/(Accessed March 15, 2021)

Drummond, R., and Wartho, R. (2009). RIMS: the Research Impact Measurement Service at the University of New South Wales. *Aust. Acad. Res. Libraries* 40, 76–87. doi:10.1080/00048623.2016.125342710.1080/00048623.2009.10721387

Gorraiz, J., Wieland, M., Ulrych, U., and Gumpenberger, C. (2020). "De Profundis: A Decade of Bibliometric Services under Scrutiny," in *Evaluative Informetrics: The Art of Metrics-Based Research Assessment*. Editors C. Daraio and W. Glanzel (Heidelberg: Springer), 233–260. doi:10.1007/978-3-030-47665-6_11

Gumpenberger, C., Wieland, M., and Gorraiz, J. (2012). Bibliometric Practices and Activities at the University of Vienna. *Libr. Manage.* 33, 174–183. doi:10.1108/01435121211217199

Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., and Rafols, I. (2015). Bibliometrics: the Leiden Manifesto for Research Metrics. *Nature* 520, 429–431. doi:10.1038/520429a

Iribarren-Maestro, I. (2018). Bibliometría Y Bibliotecas Universitarias: ¿matizando El Perfil Profesional?. *ThinKEPI* 12, 142–151. doi:10.3145/thinkepi.2018.15

Jappe, A. (2020). Professional Standards in Bibliometric Research Evaluation? A Meta-Evaluation of European Assessment Practice 2005-2019. *PloS one* 15, e0231735, doi:10.1371/journal.pone.0231735

Karolinska Institutet (2014). *Bibliometric Indicators – Definitions and Usage at Karolinska Institutet*. Available at: https://kib.ki.se/sites/default/files/bildarkiv/Dokument/bibliometric_indicators_2014.pdf (Accessed March 15, 2021)

Leiß, C. (2017). "Engaging in Performance Measurement: Introducing Bibliometric services". In 2017 IATUL Proceedings. Available at: https://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=2191&context=iatul (Accessed April 15, 2021).

Moed, H. F. (2017). *Applied Evaluative Informetrics*. Dordrecht: Springer. doi:10.1007/978-3-319-60522-7

Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Sham, M. H., Barbour, V., et al. (2020). The Hong Kong Principles for Assessing Researchers: Fostering Research Integrity. *Plos Biol.* 18, e3000737, doi:10.1371/journal.pbio.3000737

Moral-Muñoz, J. A., Herrera-Viedma, E., Santisteban-Espejo, A., and Cobo, M. J. (2020). Software Tools for Conducting Bibliometric Analysis in Science: An Up-To-Date Review. *Epi* 29, e290103. doi:10.3145/epi.2020.ene.03

NIFU (2019). *Bibliometric Analysis and Career Mapping of the SFF Scheme*. Available at: https://nifu.brage.unit.no/nifu-xmlui/bitstream/handle/11250/2655090/NIFUreport2019-31.pdf (Accessed March 19, 2021)

Pacheco-Mendoza, J., Alhuay-Quispe, J., and Machin-Mastromatteo, J. D. (2020). Bibliometrics Units as Dynamic Engines for Universities' Scientific Production. *Inf. Develop.* 36, 301–305. doi:10.1177/0266666920918466

Petersohn, S., and Heinze, T. (2018). Professionalization of Bibliometric Research Assessment. Insights from the History of the Leiden Centre for Science and Technology Studies (CWTS). *Sci. Public Pol.* 45, 565–578. doi:10.1093/scipol/scx09210.1093/scipol/scx084

Rehn, C., and Kronman, U. (2008). *Bibliometric Indicators for Karolinska Institutet*. Available at: https://kib.ki.se/sites/default/files/bibliometric_handbook_2014.pdf (Accessed March 15, 2021) doi:10.1117/12.797252

Robinson-García, N., and Gadd, L. (2019). *Results from the 2019 Responsible Metrics State of the Art Survey*. Available at: https://thebibliomagician.wordpress.com/2019/11/12/results-from-the-2019-responsible-metrics-state-of-the-art-survey/ (Accessed April 15, 2021).

Todeschini, R., and Baccini, A. (2017). *Handbook of Bibliometric Indicators: Quantitative Tools for Studying and Evaluating Research*. Weinheim: Wiley VCH.

Torres-Salinas, D., and Jiménez-Contreras, E. (2012). Hacia las unidades de bibliometría en las universidades: modelo y funciones. *Rev. Esp. Doc. Cient.* 35, 469–480. doi:10.3989/redc.2012.3.959

Universidad de Granada (2014). *Searching Research Excellence: An In-Depth Bibliometric Analysis of the International Research Contribution of the University of Granada*. Available at: https://zenodo.org/record/1045855 (Accessed June 11, 2021)

Vinkler, P. (2010). *The Evaluation of Research by Scientometric Indicators*. Oxford: Chandos Publishing. doi:10.1533/9781780630250

Zahedi, Z., and Costas, R. (2018). General Discussion of Data Quality Challenges in Social media Metrics: Extensive Comparison of Four Major Altmetric Data Aggregators. *PloS one* 13, e0197326, doi:10.1371/journal.pone.0197326

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Research Metrics and Analytics

PERSPECTIVE
published: 06 August 2021
doi: 10.3389/frma.2021.703159

New Text

Awareness Mentality and Strategic Behavior in Science

*Rafael Ball**

Director ETH Library, ETH Zurich, Zurich, Switzerland

Acknowledgement of scientific achievements was and is essentially achieved through the citation of a publication. Increasingly, however, it is no longer just the publication itself that plays an important role, but also the degree of attention that a scientist achieves with this very publication. Thus, the importance of strategic behavior in science is progressing and an awareness mentality is spreading. In this paper, the causes and backgrounds of this development are discussed, identifying the use of reductionist, quantitative systems in science management and research funding, the loss of critical judgment and technocratic dominance, quantitative assessments used for decision making, altmetrics and the like as alternative views, the use of perception scores in reference databases and universities as well as ambitions of journals as main drivers. Besides, different forms of strategic behavior in science and the resulting consequences and impacts are being highlighted.

Keywords: scholarly communication, scientometrics, awareness, bibliometrics, academic publishing

OPEN ACCESS

Edited by:
Juan Ignacio Gorraiz, Universität Wien, Austria

Reviewed by:
Daniela De Filippo, Universidad Carlos III de Madrid, Spain
Yong Huang, Wuhan University, China

****Correspondence:***
Rafael Ball
rafael.ball@library.ethz.ch

Specialty section:
This article was submitted to Research Assessment, a section of the journal Frontiers in Research Metrics and Analytics

Received: *30 April 2021*
Accepted: *15 July 2021*
Published: *06 August 2021*

Citation:
Ball R (2021) Awareness Mentality and Strategic Behavior in Science. Front. Res. Metr. Anal. 6:703159. doi: 10.3389/frma.2021.703159

INTRODUCTION

The number of scientific publications has been increasing dramatically for decades. Every day 8,500 scientific papers are being published and annually there is an increase of three million papers evaluated in the database "Web of Science" alone, which accounts for only 5% of the scientific journal output (Oeser 1976, 121). The analysis of peer-reviewed Science and Engineering publications in the Scopus database as of July 2017 show an annual growth of 3.9% per year worldwide. This leads to a doubling of total publications from 1.5 million in 2006 to almost three million publications in 2016 (National Science Foundation 2018, 101). Large platforms such as academia.edu multiply the amount of information through the second and third publication of publications, as do the countless institutional and discipline preprint or postprint repositories (Conard, 2018, 255). The duplication of content alone because of digital dissemination possibilities leads to an ever-increasing amount of scholarly information competing for the limited attention of scholars. As a consequence of the dramatically increasing amount of information (Meadows, 1998, 15–16; Proquest, 2020) and the almost unimprovable attention capacity of human beings, there is an ever-increasing perception deficit, which in turn must be compensated for by a wide variety of awareness measures on the part of researchers. Not every publication is noticed and cited any more. A competition for attention has arisen, which is conducted by various means. These increasingly include the communication tools of social media. At the same time, proving that one's own research and output are perceived is strategically relevant, especially for young researchers, for their scientific and personal survival in the academic world in the competitive struggle for funding and positions. An awareness mentality is emerging in science paired with and at the same time as an expression of strategic behavior. As long as citation numbers and other quantitative metrics were only an end in themselves for one's own perception analysis, they could not do any major damage. Adapted behavior only became problematic and necessary in the sense of a "survival strategy" when publication figures, citation numbers, the allocation of research funding, personal careers and the professional future of academics combined to form a dangerous amalgam. "Academia has become a publish or perish world" (Moosa, 2018, 2).

This strategic behavior serves as a survival strategy in academia today. Thus, I argue that only those who behave strategically in terms of the awareness mentality have a long-term chance of survival in the academic ecosystem.

DELIMITATION AND DEFINITION OF TERMS

The "Awareness Mentality"

In the context of this publication I use the terms "awareness" and "awareness mentality" to mean attention, perception, but also awareness and knowledge or familiarity. In other words, a mentality whose guiding principle is the generation of attention, of perception and—if we break it down to the individual scientist—of the creation of his or her own perception in the sense of being perceived and made known. "Scientists trying to maximise attention must not only care about selling their product, they must also care about making it a stir" (Franck, 2002, 8). In this means, I claim—and then try to prove and explain—that today's scientist invests a large part of his or her labor in generating awareness. Or must invest. This strategic pressure to adapt leads to an inner and outer attitude that I call "awareness mentality."

Strategic Behavior in Science

A person behaves strategically when he or she is guided by internal or external goals and—since we are talking about scientific behavior—subordinates content, questions, research design, methods, and communication of results to these goals. This may be for example, be conducted by means of "salami sclicing"—e.g., publishing an excessive number of papers from a single study–as discussed by Bailey (2002). This does not correspond to the idea and principles of academic science. Scientists, in the self-referential system of science, which defines goals and questions from within itself, should not be guided by external goals that are not meant to be self-referential in the sense of science (Rheinberger, 2018). This outlines the basic topics of this article. Accordingly, it would still have to be clarified why scientists allow themselves to be guided (or have to be guided) by goals other than those intrinsic to science and by what methods and means they do so.

A SHORT LOOK AT THE HISTORY OF SCIENCE COMMUNICATION

Starting with the last question I try to deduce what may have caused the claimed awareness mentality and how it is shaping in the context of changes in scholarly communication. To do this, we take a brief look at the history of scholarly communication.

We see three paradigm shifts in scholarly communication (Ball, 2020): About 2000 years ago, there was the shift from oral to written communication, whose strongest representatives were Plato and Aristotle. After the triumph of written communication and its externalization in the most diverse forms dominated by the respective description and media techniques, the invention of letterpress printing with movable type by Johannes Gutenberg in 1,452 is the second paradigm shift in scholarly communication. It was the first time that scholarly communication became a mass topic and the dissemination of large quantities of identical texts was made possible independently of geography and time. Both the Reformation and the Renaissance of science were children of printing (Greenblatt, 2012). The third paradigm shift is the emergence of mass digitization, which radically has changed everything—from the creation of content and its production, to its dissemination and archiving.

From now on, it was possible to disseminate scientific content even faster in even greater mass and reach, but at the same time more freely and independently of the established systems of the print era. The awareness mentality picked up speed at the same time as digitality, even though the topic of awareness already had its beginnings in the pre-digital era, especially after the emergence of mass (or big) science after the end of World War II (Price and Derek, 1963).

The notion of awareness outlined here does not (or not primarily) refer to the marketing of one's own scientist persona, but to the marketing of his or her publications. This was and is the central method of drawing attention to one's own successes and qualities (Franck, 2002, 4). The currency of science has always been attention and recognition for the creation of new knowledge (Tunger, 2018). The publication and its perception created and still creates recognition. We still define four central functions of a publication today (Shorley and Jubb, 2013): registration, thanks to which the scientific findings are protected by copyright and can be cited at the same time. Certification, which proves that it is a contribution of impeccable quality (for example, through peer reviewing). Perception, which draws the attention of other scientists to one's own findings and makes them available as the subject of further research, and archiving, which guarantees long-term storage and accessibility of scientific findings for posterity.

These four functions are still fulfilled by publications in times of digitality, but today they are weighted differently and implemented through new, digital methods.

RATINGS AND RANKINGS IN SCIENCE

Acknowledgement of scientific achievements was and is essentially achieved through the citation of a publication. Through professionalization and automatic processing in citation indices, this form of awareness has led to a quantification of scientific results in which it is increasingly no longer the content of the publication that matters but the degree to which it is being perceived. This has transformed the quality assessment of content into a measurement of its perception. The run on citation and measurement metrics in science represents the beginning of an awareness mentality that subsequently has become a veritable mass movement through digital systems. The system of quantitative perception measurement is about to take on a life of its own, when successful science is no longer defined as "Being good," but as "Looking good."

Digitality enables completely new forms of production, distribution and screening of scientific content. This is the technical perspective. But also, in terms of the sociology of science, the perception of results has become an increasingly important aspect of scientific activity. Producers of scientific content are adapting to this and start acting strategically. They will and must prefer forms and formats that guarantee them the appropriate attention (Weingart, 2005, 331). These are no longer only classic journal articles, books, and conference papers. Instead, contributions in the various channels of social media and other, new, digital network formats are increasing (Weingart, 2005, 272). The change in output formats is multifactorial. New technical potentials play just as much a role as normative elements, such as the demand for open access and open science, as well as competitive pressure in career planning or the financing of positions and funding (Krull, 2017, 46).

The increasing importance of rankings and ratings of individuals, institutions and countries is used and loved by the scientific community but at the same time is being criticized and rejected. Today, there is an almost unmanageable number of indicators for evaluating publication activity and its perception (Hinz et al., 2020). This includes not only the classic citation indicators, but also increasingly alternative metrics (Haustein et al., 2014), such as those that have been collected since 2010 and, above all, depict perception in social media: "The composition of the attention score is based on an algorithm that adds up the attention of scientific output in the various sources, weighted differently" (Tunger et al., 2017, 6).

The fact alone that such perception scores are already found extensively not only on scientists' websites, but also in the established reference databases of scientific literature such as the Web of Science, Scopus or Dimensions, as well as in the repositories of universities, shows that researchers cannot avoid quantification when competing for attention and awareness.

Although "in general (...) altmetrics (should) not be seen as a substitute for classical peer review in the context of quality assessment of scientific output, but are interpreted as a way to get a second opinion and additional information" (Tunger et al., 2017, 7), the new attention scores are also and increasingly significant and must be used by researchers. It is a tightrope walk to establish helpful instruments in science and, at the same time, not to pave the way for quick incentives.

Thus, there is a tension between, on the one hand, meaningful indicators that can help researchers measure the impact of their research output. On the other hand, these same indicators put even more pressure on researchers to design their work in such a way that they achieve satisfactory values. Breaking out of this vicious circle is practically impossible, especially for young scientific careers. Quantitative assessments of publication performance are used as central decision-making aids in the allocation of funds and positions (Osterloh and Frey, 2015, 65).

With this practice in science and publication management, we should therefore not be surprised today at the flood of publications, the rising journal and APC prices or the use of reductionist, quantitative systems in science management and research funding, nor at the loss of critical judgement and the marching through of technocrats (Andersson, 2008, 16).

CAUSES OF STRATEGIC BEHAVIOR IN SCIENCE

The causes for strategic (mis) behavior and thus for the attempt to gain more attention and more citations are manifold, even if they have not yet been researched in detail (Huberts, 2014).

Essentially, the causes are to be found in the pressure of the overall system of science, including its culture of communication and publication. Certainly, questions of personal responsibility and morality also play a role since there are also people among scientists who do not take the truth very seriously. The coupling of the number and quality of publications (and thus perceptions) with the awarding of funding and career options is symptomatic and puts scientists under enormous pressure (Hall and Martin, 2019, 414). The system sets false incentives and causes despair, fear of losing one's job and livelihood, career crumbling, and the fear of not being awarded funding.

But the expectations of scientific journals are also increasingly rising. In the last 10 years, the number of rejected article submissions has increased tenfold (Hall and Martin, 2019, 414). Under pressure from editors (who in turn are under pressure from publishers), journals must publish ever more spectacular findings and reports. Competition among journals and the assessment of their importance and quality through performance indicators and awareness also lead to fierce competition in the publishing industry. This is passed on to the authors and puts additional pressure on scientists.

False incentives in the reward system of science cause extremely high expectations of results in Asia, for example. In particular, the expectation of the social elites concerning visible success of science in their countries is the cause of enormous personal pressure on researchers and one of the causes of questionable publications from Asia (Lee and Schrank, 2010).

VARIOUS FORMS OF STRATEGIC BEHAVIOR IN SCIENCE

Strategic behavior in science can take many different forms: You can subordinate research to trendy topics and thus try to obtain as much funding as possible, you can carry out spectacular experiments with the attempt to achieve spectacular results. Or you can optimize your citation rate through the strategic selection of the publication media. "Negative results," failed experiments or hypotheses that cannot be confirmed, on the other hand, are neither desired in the scientific publication system, nor do they achieve the necessary acceptance and attention, nor do they survive the peer review process (ScienceMatters, 2021).

None of this constitutes scientific misconduct. In the sense of increasing success and optimizing awareness, it is morally questionable in the worst case, but not reprehensible (if the category of morality may be applied here at all). Nevertheless, the strategy of achieving (or having to achieve) scientific success through increased awareness leads, in the age of digitality, to a situation where attention is no longer to be achieved solely through the classic citation metrics, but increasingly through the systems of social media, which operate in a short-lived, fast, high-frequency, and non-stop manner (Barth, 2019, 9). Thus, the system requires scientists today not only to have in-depth knowledge of social media and their use, but also to adequately prepare appropriate content, which differs significantly from classic publications, especially in the scope and depth of the argumentation (Ram and Rameshwar, 2016, 229).

It is only a very small step from a purely objective, reserved, research, and knowledge-driven science to a working manner that formulates questions, designs research experiments, and presents results in such a way that it primarily serves to generate attention. However, not everyone who uses social media skillfully and thus makes his or her research better known than through pure "citation perception" is conducting strategically controlled or even dubious research.

Strategic action, scientific misconduct, and manipulation move close together in a narrow corridor. The more pressure there is on researchers to generate attention for their results, the faster the distinction between the two becomes blurred. If "Looking good" is better rewarded than "Being good," the inhibition threshold for the step from knowledge-driven science to an attention-seeking show decreases. And with it the quality of research, whose results are increasingly no longer reproducible, which causes the already discussed "crisis of reproducibility," or at least exacerbates it (Moosa, 2018, 71–73).

Questionable and inappropriate behavior in science is becoming increasingly common. In a survey conducted by Bouter (2015), 43% of all researchers admitted to have published questionable data and results, 2% even to have deliberately falsified. When asked about their opinion of other scientists, it was suspected that 14% of others falsify and 72% publish questionable results.

From careless editing of data to deliberate fraud, all forms of manipulation are being demonstrated. Boundaries are being extended, hypotheses are being adapted to the results and vice versa. Methods are being falsified or rearranged, as is the underlying data. Other phenomena observed are text recycling, self-plagiarism and genuine plagiarism (Öchsner, 2013, 95).

It is reasonable to assume that at least a certain proportion of such wrongdoing can be traced back to an exaggerated or mislead form of awareness mentality. Against this background the inappropriate behavior in science can be understood as an expression of the aim to gain attention and secure livelihoods.

In one study a significant increase in titles of scientific publications in medicine, life sciences and physics that end with a question mark is shown (Ball, 2007). For example, in medicine, the proportion of "question mark articles" increased eightfold in the study period. A random qualitative analysis showed that scientists increasingly chose daring or spectacular titles for their publications to attract attention. To nevertheless remain scientifically credible, the titles are ended with a question mark, so that there is still the option of retraction–just in case.

This in turn leads to some challenges, particularly related to peer reviewing. On the one hand, scientific misconduct can simply not always be identified in a peer review process—especially not when we are in the gray area between strategic behavior to attract attention and actual misconduct. This reduces the assurance of certification, which as one of the four basic functions of a publication mentioned above should also be achieved through peer reviewing.

Another challenge of peer reviewing is the already mentioned fact that negative results often do not survive the peer reviewing process. If this does not change in the future, it is to be feared that peer reviewing will fuel a misunderstood and exaggerated form of awareness mentality.

IMPACTS AND CONSEQUENCES

The four basic functions of a publication (Shorley and Jubb, 2013) have not become meaningless even in the age of digitality, social media and awareness hype. The individual scientist must still fulfil the basic functions of a publication. However, it is no longer enough to send the manuscript to a publisher and to wait for the paper to be accepted: This is only the basic step of a publication practice that has changed fundamentally. Today, proof must also be provided that the paper is available in Open Access. Today, depositing the paper in an institutional or subject-specific repository is just as much an obligation as supplementing the article with the research data and references used, as well as linking it in various academic networks. Blog posts about new findings of the paper are expected, as is the use of Twitter and other social media to draw attention to the new paper. Marketing for the purpose of generating awareness has an increasing share in the publication effort. In addition, the pure academic community and the interested public are increasingly merging. If you really want to attract attention to yourself and your research, another need is to present and explain your results in a generally understandable way. In doing so, the results not only serve their own specialist community, but also a general audience. Video messages and Instagram appearances complete the awareness campaign.

In a possible perspective of scientific publishing, we have to state that the classical concept of publication with all its implications will dissolve if the success of a publication is no longer measured only by the truthfulness of the messages, but by the mere determination and quantification of the perception of what is communicated (and no longer the perception itself). If perception (and its ascertainment) becomes to be more important than truth, then all barriers will fall for the uncontradicted boundary shift from knowledge to opinion and vice versa. "Looking Good" becomes more important than "Being Good," also because a rapidly increasing number of publications makes qualified reception impossible. If scientific results are then increasingly no longer reproducible, the common sense of the basic principle of publishing scientific findings for the purpose of their reception, discussion and further development will finally go bankrupt.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

REFERENCES

Andersson, K. 2008. Transparency and Accountability in Science and Politics. *Transparency and Accountability in Science and Politics: The Awareness Principle. 2008*. Edition. Houndmills, Basingstoke, Hampshire; New York: Palgrave Macmillan. doi:10.1057/9780230227767

Bailey, B. J. (2002). Duplicate Publication in the Field of Otolaryngology-Head and Neck Surgery. *Otolaryngol. Head Neck Surg.* 126 (3), 211–216. doi:10.1067/mhn.2002.122698

Ball, R. (2007). "Wissenschaftskommunikation im Wandel: Die Verwendung von Fragezeichen im Titel von wissenschaftlichen Zeitschriftenbeiträgen in der Medizin, den Lebenswissenschaften und in der Physik von 1966 bis 2005". *Inf. - Wissenschaft Praxis* 58 (6–7), 371–375. doi:10.3929/ethz-b-000193627

Ball, R. (2020) in *Wissenschaftskommunikation im Wandel: Von Gutenberg bis Open Science* (. VS Verlag für Sozialwissenschaften). doi:10.1007/978-3-658-31541-2

Barth, Alfred. (2019). *Publish or Perish!: Ein Schwarzbuch Der Wissenschaft*. Nordhausen: Traugott Bautz.

Bouter, L. M. (2015). Commentary: Perverse Incentives or Rotten Apples? *Account. Res.* 22 (3), 148–161. doi:10.1080/08989621.2014.950253

Conard, J. (2018), Die fatale Folge von Sensationsmeldungen. *Fake oder Fakt? Wissenschaft, Wahrheit und Vertrauen, herausgegeben von Carsten Könneker*, Berlin, Heidelberg: Springer, 251–258. doi:10.1007/978-3-662-56316-8_22

Franck, G. (2002). "The Scientific Economy of Attention: A Novel Approach to the Collective Rationality of Science". *Scientometrics* 55 (1), 3–26. doi:10.1023/A:1016059402618

Greenblatt, S. (2012). *Die Wende: wie die Renaissance begann. Aufl. München: Siedler*. 6. doi:10.1163/25890581-089-01-90000018

Hall, J., and Martin, B. R. (2019). Towards a Taxonomy of Research Misconduct: The Case of Business School Research. *Res. Pol.* 48 (2), 414–427. doi:10.1016/j.respol.2018.03.006

Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H., and Terliesner, J. (2014). Coverage and Adoption of Altmetrics Sources in the Bibliometric Community. *Scientometrics* 101 (2), 1145–1163. doi:10.1007/s11192-013-1221-3

Hinz, O., van der Aalst, W. M. P., Weinhardt, C., and Weinhardt, Christof. (2020). Research in the Attention Economy. *Bus Inf. Syst. Eng.* 62 (2), 83–85. doi:10.1007/s12599-020-00631-6

Huberts, L. (2014), The Integrity of Governance. *The Integrity of Governance: What it Is, what We Know, what Is Done and where to Go. 2014*. Edition. Basingstoke; New York: Palgrave Macmillan. doi:10.1057/9781137380814

Krull, Willhelm. (2017). *Die vermessene Universität: Ziel, Wunsch und Wirklichkeit. Herausgegeben von Günther Burkert, Wilhelm Krull, Antonio Loprieno, und Eva Barlösius. 1*. Edition. Wien: Passagen.

Lee, C.-S., and Schrank, A. (2010). Incubating Innovation or Cultivating Corruption? the Developmental State and the Life Sciences in Asia. *Social Forces* 88 (3), 1231–1255. doi:10.1353/sof.0.0282

Meadows, A. J. (1998). Communicating Research. *Library and Information Science*. New York: Academic Press.

Moosa, Imad. A. (2018). *Moosa, I: Publish or Perish: Perceived Benefits versus Unintended Consequences*. Cheltenham, UK Northampton, MA: Edward Elgar Publishing. doi:10.4337/9781786434937

National Science Foundation (2018). "Science and Engineering Indicators 2018". Available at: https://www.nsf.gov/statistics/2018/nsb20181/report.

Öchsner, A. (2013), Introduction to Scientific Publishing. *SpringerBriefs in Applied Sciences and Technology*. Berlin Heidelberg: Springer-Verlag. doi:10.1007/978-3-642-38646-6

Oeser, Erhard. (1976). Wissenschaft und Information sytematische Grundlagen einer Theorie der Wissenschaftsentwicklung Bd. 1 Wissenchaftstheorie und empirische Wissenschaftsforschung. *Oeser, Erhard Wissenschaft und Information 1 (RERO)0552544-41slsp*. Wien: ROldenbourg. doi:10.7767/boehlau.9783205159599

Osterloh, M., and Frey, B. S. (2015). Rankings und der Preis der Wissenschaft. *Z. für Kulturwissenschaften* 9 (1), 65–78. doi:10.14361/zfk-2015-0110

Price, S., and Derek, J. de. (1963). *Little Science, Big Science*. New York: Columbia University Press. doi:10.7312/pric91844

Proquest (2020). O. J. "Ulrich's Periodicals Directory™ (57th Edition) 2019 - Ulrich's Periodicals Directory™ (57th Edition) 2019". Zugegriffen 23. Available at: https://www.proquest.com/products-services/related/Ulrichs-Periodicals-Directory.html (April, 2020).

Ram, Shri., and Rameshwar, Rudra. (2016). "Measuring Research Impact of Scholarly Publications Using Altmetrics", Scholarly Communication and the Publish or Perish Pressures of Academia*, von Achala Munigal, 1*. Edition. Hershey PA: IGI Global. doi:10.4018/978-1-5225-1697-2

Rheinberger, H.-J. (2018). *Experimentalität: Hans-Jörg Rheinberger im Gespräch über Labor, Atelier und Archiv*. Berlin: Kadmos.

ScienceMatters (2021). o. J. Zugegriffen 20. Available at: https://www.sciencematters.io/ April, 2021).

Shorley, Deborah, Michael Jubb (eds.). (2013). *The Future of Scholarly Communication*. London: Facet. https://doi.org/10.29085/9781856049610.

Tunger, Dirk., Meier, Andreas., and Hartmann, Hans-Daniel. (2017). Machbarkeitsstudie Altmetrics". FZJ-2018-00250. Zentralbibliothek. Available at: https://juser.fz-juelich.de/record/841963.

Tunger, Dirk. (2018). "Altmetrics: Kommt die, Ökonomie der Aufmerksamkeit'?" *Wissenschaftskommunikation.de* (blog). Available at: https://www.wissenschaftskommunikation.de/altmetrics-kommt-die-oekonomie-der-aufmerksamkeit-21609/ (November 28, 2018). doi:10.5772/intechopen.76874

Weingart, Peter. (2005). "Die Stunde der Wahrheit? Studienausgabe: Zum Verhältnis der Wissenschaft zu Politik," in *Wirtschaft, Medien in der Wissensgesellschaft. 2. Auflage 2005 Edition* (Weilerswist: Velbrück Wissenschaft).

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership