

Stock, Wolfgang G.; Reichmann, Gerhard; Dorsch, Isabelle; Schlögl, Christian

Article — Published Version

Counting Research Publications, Citations, and Topics: A Critical Assessment of the Empirical Basis of Scientometrics and Research Evaluation

Journal of Information Science Theory and Practice

Suggested Citation: Stock, Wolfgang G.; Reichmann, Gerhard; Dorsch, Isabelle; Schlögl, Christian (2023) : Counting Research Publications, Citations, and Topics: A Critical Assessment of the Empirical Basis of Scientometrics and Research Evaluation, Journal of Information Science Theory and Practice, ISSN 2287-4577, Korea Inst. of Science and Technology Information, Daejeon, Vol. 11, Iss. 2, pp. 37-66,
<https://doi.org/10.1633/JISTaP.2023.11.2.4>

This Version is available at:

<http://hdl.handle.net/11108/577>

Kontakt/Contact

ZBW – Leibniz-Informationszentrum Wirtschaft/Leibniz Information Centre for Economics
Düsternbrooker Weg 120
24105 Kiel (Germany)
E-Mail: info@zbw.eu
<https://www.zbw.eu/de/ueber-uns/profil-der-zbw/veroeffentlichungen-zbw>

Standard-Nutzungsbedingungen:

Dieses Dokument darf zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden. Sie dürfen dieses Dokument nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Sofern für das Dokument eine Open-Content-Lizenz verwendet wurde, so gelten abweichend von diesen Nutzungsbedingungen die in der Lizenz gewährten Nutzungsrechte.

Terms of use:

This document may be saved and copied for your personal and scholarly purposes. You are not to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public. If the document is made available under a Creative Commons Licence you may exercise further usage rights as specified in the licence.



<http://creativecommons.org/licenses/by/4.0/>

Counting Research Publications, Citations, and Topics: A Critical Assessment of the Empirical Basis of Scientometrics and Research Evaluation

Wolfgang G. Stock* 

Department of Information Science, Heinrich Heine University
Düsseldorf, Düsseldorf, Germany
E-mail: wolfgang.stock@hhu.de

Isabelle Dorsch 

ZBW – Leibniz Information Centre for Economics, Kiel, Germany
E-mail: i.dorsch@zbw-online.eu

Gerhard Reichmann 

Department of Operations and Information Systems, University of Graz,
Graz, Austria
E-mail: gerhard.reichmann@uni-graz.at

Christian Schlägl 

Department of Operations and Information Systems, University of Graz,
Graz, Austria
E-mail: christian.schloegl@uni-graz.at

ABSTRACT

Scientometrics and research evaluation describe and analyze research publications when conducting publication, citation, and topic analyses. However, what exactly is a (scientific, academic, scholarly or research) publication? This article demonstrates that there are many problems when it comes to looking in detail at quantitative publication analyses, citation analyses, altmetric analyses, and topic analyses. When is a document a publication and when is it not? We discuss authorship and contribution, formally and informally published documents, as well as documents in between (preprints, research data) and the characteristics of references, citations, and topics. What is a research publication? Is there a commonly accepted criterion for distinguishing between research and non-research? How complete and unbiased are data sources for research publications and sources for altmetrics? What is one research publication? What is the unit of a publication that causes us to count it as “1”? In this regard, we report problems related to multi-author publications and their counting, weighted document types, the unit and weighting of citations and references, the unit of topics, and counting problems—not only at the article and individual researcher level (micro-level), but also at the meso-level (e.g., institutions) and macro-level (e.g., countries). Our results suggest that scientometric counting units are not reliable and clear. Many scientometric and research evaluation studies must therefore be used with the utmost caution.

Keywords: scientometrics, research publication, counting unit, reference, citation, topic

Received: October 12, 2022
Accepted: February 21, 2023

Revised: February 18, 2023
Published: June 30, 2023

***Corresponding Author:** Wolfgang G. Stock
 <https://orcid.org/0000-0003-2697-3225>
E-mail: wolfgang.stock@hhu.de



All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee and authors' permission. All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). Under this license, authors reserve the copyright for their content; however, they permit anyone to unrestrictedly use, distribute, and reproduce the content in any medium as far as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

1. INTRODUCTION

Scientometrics is a research area which focuses on formal scholarly communications (Borgman & Furner, 2002; Leydesdorff, 1995) and on likewise informal communications with the research environment, which include internal informal communication amongst researchers and external communication between researchers and the general public. “Communication” means that there are sustainable documents such as journal articles, papers in proceedings, patents, books, and postings in social media services, but not ephemeral documents such as oral communications. Formal scholarly communication happens via research documents, which deal with research topics and nearly always contain references. Informal communications with the research environment comprise all kinds of mentions in social media, in the newspapers, or in other mass media, some of them studied by altmetrics (Thelwall et al., 2013). Scientometric results play a role in the evaluation of research units at different levels in relation to single researchers, institutions, journals, disciplines, or even countries, in the analysis of age- and gender-related science studies (Larivière et al., 2013; Ni et al., 2021) as well as in research policy and management (Mingers & Leydesdorff, 2015), too. We use the term “scientometrics” as a narrower term of “informetrics.” While scientometrics studies research information, “informetrics” is the study of all quantitative aspects of all kinds of information (Dorsch & Hausteine, 2023; Stock & Weber, 2006; Tague-Sutcliffe, 1992).

According to Leydesdorff (1995, p. 4), the science of science in general has three dimensions: scientists, cognitions, and texts. The subjects of scientometrics are the multiple relationships between scientists and their texts, i.e., the published research information. The remaining disciplines of the science of science are the sociology of knowledge, studying the relationships between scientists and cognitions, and the theories of information and communication, which analyze the relationships between cognitions and texts. In scientometrics, the “methodological priority of textual data” dominates (Leydesdorff, 1995, p. 57).

Scientometrics includes quantitative studies on all kinds of research. Particularly, quantitative research studies call for a solid empirical observation base. Each research field has its own technical terms which directly relate to observation objects and their properties, and sentences reflecting protocols of propositions with those technical terms (Carnap, 1931; 1932; Neurath, 1932).

Important elementary concepts of scientometrics are “research publication” and “word in a research document,” or, alternatively, “index term” in a record of a bibliographic information service, “reference” / “citation,” and “mention in the research environment.” A protocol sentence may be “X has 350 publications as of 2023-01-01.” Scientometricians *measure* counts of publications, words, index terms, references, citations, and mentions, as these objects are directly observable. In scientometrics, numbers of publications are an *indicator* for research output, words in research documents or index terms in bibliographic records are an indicator for research topics, and numbers of citations are an indicator for research impact, but not for research quality, as this is not measurable at all (Glänzel, 2008); and, finally, in altmetrics, the numbers of mentions in the research environment, for instance in research blogs, in microblogs, in social media services, or in newspapers, can be interpreted as an indicator of research interestingness or attention the research receives. A sentence such as “X’s research productivity is higher than average” would be a typical proposition applying a research indicator. We should have in mind that research output or research productivity, which is output in relation to input (Stock et al., 2023), topics, impact, and interestingness are not directly observable but only mediated by measures and deviations from measures, for instance, publications per year or co-citations.

References are hints for information flows from previous research towards a research publication; citations are hints for information flows towards further research; however, citations are references in the citing publications. As research publications contain words and references and so indirectly also citations, an important empirical observation base of scientometrics is publications. Here an issue arises: “What is a scientific, academic, scholarly, or research publication?” This issue leads to three research questions (Reichmann, 2020; Stock, 2000; 2001):

1. What is a publication? When is a research document a publication and when is it not? What are references and citations? And what is a topic?
2. What is a research publication? Are there criteria for demarcation of research from non-research? What are appropriate data sources for research publications?
3. What is one research publication? What is the unit of a publication inducing us to count it as “1?” And, of course, additionally, what are *one* reference, one citation, and one topic?

2. METHODS

This paper intends to answer these research questions by means of a critical review of the state of the art in scientometrics and research evaluation, i.e., we collected literature and tried to generate new knowledge about the nature of the properties of research publications as the base unit for many scientometric studies.

The reviewed literature was found through searches in bibliographic information services (Web of Science [WoS], Scopus, Google Scholar, and Dimensions) and on SpringerLink, as Springer is the publishing house of the journal *Scientometrics*, applying both the citation pearls growing strategy and the building block strategy (Stock & Stock, 2013, pp. 261-263) wherever possible. Additionally, we followed the links backwards via references and forwards via citations of found articles as a kind of snowball sampling. We collected relevant hits from the direct database results and from their references and citations; additionally, we found some hits from browsing and serendipity, thus applying a berrypicking strategy (Bates, 1989) in the end.

We searched separately for every topic, as a global search strategy failed. For, e.g., “criterion of demarcation” we initially looked for “demarcation criterion” and hoped to find some very relevant “citation pearls.” In the next step, we checked the found pearls for further search arguments such as, for instance, “pseudo-science,” “science and non-science,” or “problem of demarcation” and put them together in the sense of the building block strategy to: “(science OR ‘pseudo-science’ OR ‘non-science’) AND (‘demarcation criterion’ OR ‘criterion of demarcation’ OR ‘problem of demarcation’).” In the last step, we browsed through the references and citations.

This paper is not a scientometric research study and, in particular, it does not provide a quantitative analysis. It is a paper on fundamental *problems of scientometrics*, i.e., on problems of the quantitative research approach in the science of science. Our focus is on the main research object of scientometrics: the research publication, including its authors, topics, references, citations, and mentions in social media sources, but also the associated questions and problems. Are quantitative scientometric studies really free from problems with the correct collection of empirical data, the use of counting methods and indicators, and methodological problems? If there were any fundamental problems with scientometrics, would its application in research evaluation really be helpful? Or is it dangerous to rely solely on scientometric results in research description, analysis, evaluation, and policy?

It is not the purpose of this study to mention all relevant literature, but to draw a comprehensive picture of research publications as counting units of scientometrics and—first and foremost—on all related problems. The main contribution of our paper is a critical assessment of the empirical foundations of scientometrics: Does scientometrics really have empirical evidence? Is scientometrics a sound scientific discipline? Or are there so many problems with data, indicators, and research practices that scientometrics and research evaluation as scientific disciplines should be fundamentally reconsidered?

We can reveal one of this study’s main results already now: There are many serious problems. Many scientometricians are aware of these issues, but researchers from other knowledge domains conducting scientometric studies may not be. Scientometric publications—conducted by scientometric laypeople—are easy to find in many research areas. Therefore, this article is not only addressed to students and early career researchers to understand important topics of scientometrics and to act very carefully with problematic data and indicators in their own scientometric research, but also to senior researchers to revisit and solve the major issues—if this is ever possible—in the evolving field of scholarly communications.

3. WHAT IS A PUBLICATION?

In this paragraph, we discuss authorship as well as formally and informally published documents in general and, more specific, published documents in research. Of course, every research publication is a publication, but not all publications are research publications. With the emergence of social media, the concept of “publication” has changed. Tokar (2012) distinguishes between publications also by and on researchers in such social media as, e.g., Facebook, Twitter, Reddit, Instagram, or YouTube and academic or scholarly publications in “classical” online as well as offline media, e.g., in journals or proceedings, documents of intellectual property, or books. Also, publications in mass media such as newspapers, TV shows, or on the radio are publications. Documents in “classical” research media are genuine formal research publications, while publications in social media and mass media are parts of the communication with the research environment and so are—from the viewpoint of the research system—informally published.

A crucial question for scientometrics is: What exactly is a research publication? And what is it not? For the decision for or against the consideration that a document is

a research publication, Dorsch et al. (2018) propose two rules, namely the rule of formal publishing and the rule of research content. Of course, there is an additional at a first glance rather self-evident norm: In the publication's byline, the name of the author is stated.

3.1. Publication and Authorship

The "self-evident" criterion of authorship is not always clear as there exists rather unfair or unethical behavior by some researchers (Flanagin et al., 1998; Teixeira da Silva & Dobránszki, 2016): There are "ghost writers," i.e., authors who are not mentioned in the byline but who contributed to the document, and there is honorary authorship, i.e., persons who did not contribute to the document but are stated in the byline. In scientometrics, however, we have no possibilities for identifying ghost writers and honorary authors in an individual document. Therefore, we can only mention this problem but are not able to show a way to solve it. With regard to honorary authorship, the number of publications per year for individual researchers could be looked at. If this number is unrealistically high, this could be an indication of the existence of honorary authorship.

In multi-authored papers, the individual authors may have different roles. The Contributor Roles Taxonomy provides a standardized set of 14 research contributions such as, e.g., conceptualization, data curation, funding acquisition, or writing the original draft, and nowadays we find many journals indicating those contributor roles (Larivière et al., 2021). Of course, we have to presuppose that the given information is correct and unbiased. Taken seriously, we have to shift away from authorship and to move to contributorship instead (Holcombe, 2019).

3.2. Formally and Informally Published Documents

Authors publish documents or contribute to documents. But what is a document? In a first approximation, we follow Briet (1951/2006, p. 10) in defining a "document" as "any concrete or symbolic indexical sign [*indice*], preserved or recorded toward the ends of representing, of reconstituting, or of providing a physical or intellectual phenomenon". It is intuitively clear that all manners of texts are documents. But for Briet, non-textual objects may also be documents. The antelope in a zoological garden she described has become the paradigm for the additional processing of non-textual documents in documentation. An antelope in a zoo is a document and has nothing in common with a publication, but an artifact of engineering, say, a bridge or a software program, does

(Heumüller et al., 2020; Jansz & le Pair, 1991). Following Buckland (1997), objects are "documents" if they meet the following four criteria: materiality: they are physically—including the digital form—present; intentionality: there is a motivation behind them, and they carry meaning; development: they are created; and perception: they are described and perceived as documents. This broad definition of a "document" includes textual and non-textual, and digital and non-digital resources. Every text, every data set, every "piece" of multimedia on the web, every artwork, and every engineering artifact is a document.

What is a formally published document? These are all documents from publishing houses, i.e., books, newspapers as well as journals, all intellectual property documents, grey literature, which are publications not provided by commercial publishers like university publications, working papers, or company newspapers (Auger, 1998), and all audio-visual materials from broadcasting companies.

All documents which are not publicly accessible have to be excluded. Such documents are, for instance, letters and other types of personal correspondence unless the correspondence is published, and confidential documents from companies and institutions including expert opinions and technical reports to be kept under lock and key.

Many web documents including all social media documents with user-generated content are informally published. Here we find the empirical basis for many studies in altmetrics. In "outbreak science" (Johansson et al., 2018), especially during the Coronavirus disease 2019 (COVID-19) pandemic, many research papers, including preprints, were described and discussed both in social media and newspapers (Massarani & Neves, 2022) and the general amount of preprints for both COVID-19 and non-COVID-19 related papers seemed to have increased (Aviv-Reuven & Rosenfeld, 2021). During COVID-19, the role of preprints further advanced and so the science communication landscape changed as well (Fraser et al., 2021; Watson, 2022). The status of preprints in the research system seems to alter. Since 2013, the number of preprint server is increasing (Chiarelli et al., 2019; Tennant et al., 2018). Accordingly, the following questions arise: What is the position of preprints? To which extent can they be considered as research publication? Based on our definition of formal and informally published documents, they would be positioned somewhere in-between, as they in most cases might undergo peer review. In general, there exist different definitions of preprints within the literature, concerning their genre, position over time, versioning, ac-

cessibility, responsibility, and value (Chiarelli et al., 2019). However, it is always possible that an author publishes any kind of paper with any kind of content on a preprint server without any quality control (Añazco et al., 2021).

Also, research data are provided by specialized services, which were simply uploaded. However, there are formally published documents which are based on the data. Hence, both preprints and research data can be considered as documents in between formal and informal publishing.

Most publications are written in a discipline-specific publication culture. This “culture” includes the preference for different document types, for example, journal articles in the natural sciences, proceedings papers in computer science, or books in the humanities, as well as differences in the mean length of reference lists and in the mean number of citations received per paper. Additionally, some document types such as, for instance, review articles tend to get more citations than others. Field-specific publication and citation cultures, different document types, but also field-specific differences in informally published sources require normalization in order to make figures from various research disciplines, document types, or social media mentions comparable (Haustein, 2012, pp. 274-289). Normalization is usually based on the respective mean values of a field, document type, and social media service or on a combination of the three aspects. An important problem is the granularity of the definition of the field as it may be defined at various levels, from small research areas to broad academic disciplines (Zitt et al., 2005).

Following library science, a published document is a unit comprising two areas: intellectual or artistic content and physical (also digital) entities, whereby the terms “work” and “expression” are related to the content and “manifestation” and “item” to the physical entity: “The entities defined as *work* (a distinct intellectual or artistic creation) and *expression* (the intellectual or artistic realization of a *work*) reflect intellectual or artistic content. The entities defined as *manifestation* (the physical embodiment of an *expression* of a *work*) and *item* (a single exemplar of a *manifestation*), on the other hand, reflect physical form” (IFLA Study Group on the Functional Requirements for Bibliographic Records, 1998, p. 12). Works, i.e., authors’ creations, can thus have one or more forms of expression (concrete realizations, including translations or illustrations), which are physically reflected in a manifestation such as, e.g., one printing of a book, and the single items of this manifestation. For scientometrics, it makes a difference to count works, expressions (say, a work in

English and the same in German), or manifestations (the same expression in different journals or books or in different editions). Even the items are of interest for scientometrics, be it as the number of printed copies or downloads of a book edition, the circulation of a journal, or the download numbers of an article. In cataloging, there is a cut-off point between the same work such as, e.g., a translation or an edition in multiple sources, and a new work, for instance, the work’s summary or a free translation (Tillett, 2001). There are some special problems concerning manifestations. One problem is concerned with gen-identical documents, i.e., documents with slightly changing content over time. These documents can be found in loose-leaf binders such as with many legal documents, or on websites. How can we count manifestations with several parts? And how can we count documents with shared parts, e.g., the same text body but a new title (Stock & Stock, 2013, pp. 572-573), or even the same document in multiple manifestations (Bar-Ilan, 2006)? Should such duplicate or redundant publications (Lundberg, 1993) be skipped in scientometric analyses? A related problem is the “salami slicing” of research results into “least publishable units” (Broad, 1981). If such units are redundant and aiming for a quick gain of publication and citation numbers, it is indeed an issue for the research system and for scientometrics as well. However, different small salami slices of the same research project can be a beneficial mechanism in the research system when it facilitates data management and rapid publication, among others (Refinetti, 1990), of course, always under the premise of being truly necessary. Additionally, some journals strongly encourage short papers (Roth, 1981). An open question is how to decide whether a paper is small but original, or redundant or even a duplicate.

The scientometric literature on works, expressions, etc. and on non-textual documents is rather limited. What we additionally miss are precise descriptions of the concepts of “formally published documents,” “informally published documents,” and “documents in between” like preprints or research data, as the units of many or even of all scientometric studies.

3.3. What is a Citation? What is a Reference?

Almost always publications do not stand in isolation, but have relations to other publications. With regard to textual publications in contrast to works of tonal art, pictures, or movies this is described as “intertextuality” (Rauter, 2005). Intertextuality is expressed by quotations or references. Important for scientometrics are the refer-

ences, which are bibliographic descriptions of publications authors have read and used while writing their papers. References are set by the authors themselves, or—for example, in patents—by others (in patents, by patent examiners). References can be found in the text, in footnotes, end notes, in a list of prior art documents in patents, or in a bibliography. References thus show information flows towards a publication or concerning patents towards a patent examination process. At the same time, we can see in the reverse direction how reputation is awarded to the cited document. From the citer's perspective, an end note, an entry in the bibliography, etc. is a reference, whereas from the citee's perspective, it is a citation.

The act of setting references is not unproblematic. It has been known for decades that authors do not cite all influences, that there is unspecific self-citing, and that some references are meant negatively (MacRoberts & MacRoberts, 1989), including pejorative, disagreeing, condemning, and other negative denotations. Additionally, some researchers concert reciprocal citations, thus forming a citation cartel (Poppema, 2020). For MacRoberts and MacRoberts (2018), citation analysis is a mismeasure of science. The starting position for both reference analysis and citation analysis is also bad, since one cannot trust the authors as to whether they actually cite everything they have read and used and whether they do not cite anything else.

If citations are used as an indicator for research impact, it should also be mentioned that there is a difference between research and practitioner journals. The former usually have an orientation on other researchers who write articles and potentially cite the read papers. The latter have an orientation on practitioners who also read, but do not always write their own papers and therefore cannot cite the read material so often. As a consequence, practitioners are less regarded in citation analyses (Scanlan, 1988; Schlögl & Stock, 2008).

3.4. What is a Topic?

Publications are about something; they are dealing with topics. Topics of research publications must be identified so that they can be captured in scientometric topic analyses (Stock, 1990). In information science, we speak of a publication's "aboutness" (Maron, 1977; Stock & Stock, 2013, pp. 519-523). Ingwersen (2002) differentiates between indexer aboutness (the indexer's interpretation of the aboutness in a bibliographic record) and author aboutness (the content as it is) among others. It is possible but very difficult to describe the author aboutness in a quan-

titative way; here, scientometricians have to work with the words (including the references) in the publication and have to prepare more informative indicators such as, for instance, word counts or co-word analyses. It is also a possible path of topic analysis to restrict to the publications' titles and authors' keywords, if they really reflect all discussed topics. Furthermore, scientometricians can work with citation indexing, as a citation, for Garfield (1979, p. 3), may represent "a precise, unambiguous representation of a subject."

The easier way is to work with the indexer aboutness, i.e., the topics in their representations through knowledge organization methods (Haustein, 2012, p. 78), i.e., nomenclatures, classification systems, thesauri, and the text-word method. In this way, the aboutness is captured by index terms. If we apply knowledge representation systems for scientometric topic analysis, we have to deal with the respective method (e.g., classifications have notations which are built without natural languages, while thesauri work with natural language descriptors), with the concrete tool (e.g., the International Patent Classification with all its rules or the Medical Subject Headings [MeSH] with their sets of descriptors and qualifiers), and, finally, with the tool's indexing rules and the concrete work of an indexer.

4. WHAT IS A RESEARCH PUBLICATION?

"Research" covers all activities in the natural sciences, the engineering sciences, the social sciences, and the humanities. However, depending on the sources, the definitions of research differ. The Organisation for Economic Co-operation and Development (OECD) (2015), for instance, defines research very broadly in terms of progress and new applications, while, according to Popper (1962), it is necessary for scientific results to be falsifiable. In this section, we describe the criterion of demarcation of science in order to find a boundary between research and non-research. Then we discuss data sources of research publications including bibliographic information services as well as personal and institutional repositories. Finally, we describe advantages and shortcomings of altmetrics, which analyze research-related publications in social media.

4.1. Criterion of Demarcation

In the philosophy of science, authors discuss criteria for demarcation between science and non-science or pseudo-science. For Carnap (1928, p. 255), *Wissenschaft* formulates questions which are answerable in principle.

The answers are formed through scientific concepts, and the propositions are basically true or false. Please note that the German term *Wissenschaft* includes natural sciences, social sciences, and humanities. Scientific sentences are protocol sentences reflecting observable facts; Carnap's (1931) prototype of a scientific language is the language of physics. Following Neurath (1932) such a protocol sentence is defined through justifications. If a justification fails, the sentence can be deleted. Therefore, *Wissenschaft* is always on the go. Popper (1935, p. 15) reformulates Neurath's definition: An empirical-scientific system must be able to fail because of experience; so, Popper (1962) calls only falsifiable, testable, or refutable propositions scientific. However, for Kuhn (1970) there are problems when we pay attention to scientific development. Sometimes there is a radical change in the development of a discipline, a so-called "scientific revolution," and we meet supporters of the old view and at the same time supporters of the new one. For both groups, believing in their own "normal science," led by an accepted "paradigm," the other group performs non-science and they themselves do good science. Lakatos (1974) calls Popper's attempt "naïve," but also declines Kuhn's theory of science, as there is only a demarcation criterion from the view of a special normal science and no general criterion; in fact, we cannot speak about progress in science, if we agree with Kuhn's approach (Fasce, 2017). If we—with Feyerabend (1975)—accept that in science "anything goes," then there will be no demarcation criterion at all. For Lakatos (1974), progressive research programs outperform degenerating ones. A research program is progressive if its theory leads to the discovery of hitherto unknown facts. Now, progress is the demarcation criterion; progress leading to new applications, too (Quay, 1974). Also, for OECD (2015)'s Frascati Handbook, new knowledge and new applications determine science. According to the OECD (2015, p. 45), research and development activities in natural sciences, social sciences, humanities and the arts, and engineering sciences must be novel, creative, uncertain, systematic, and transferable and/or reproducible.

Fernandez-Beanato (2020, p. 375) sums up: "The problem of demarcating science from nonscience remains unsolved." As all empirical sciences and also all formal sciences (e.g., mathematics) can be falsified, they are science for Popper (1935); and religion (Raman, 2001), metaphysics, astrology (Thagard, 1978), art, or fiction books are not. Following the philosophical ideas of Popper (1935; 1962), Carnap (1928; 1931; 1932), and Neurath (1932), psychoanalysis is pseudo-science, because there is no

empirically observable data base; for Grünbaum (1977), it is science as it is successfully applicable. Parapsychology may be falsified, so it is science (Mousseau, 2003); so is—against common sense—homeopathy (Fasce, 2017, p. 463). The philosophy of science clearly shows the problems of demarcation. But it has no satisfying answer for the use of scientometrics. Perhaps it is better to formulate acceptable criteria for pseudo-science (e.g., Fasce, 2017, p. 476 or Hansson, 2013, pp. 70-71). These are Fasce's (2017) criteria: Pseudo-science is presented as scientific knowledge, but it refers to entities outside the domain of science, applies a deficient methodology, or is not supported by evidence. However, this definition makes use of the here undefined terms of science, scientific knowledge, and evidence. If we are going to define these concepts we must start from the beginning.

In scientometrics, it would be very problematic to check demarcation for every single research publication. Since scientometric studies often deal with large amounts of data, such a check would not even be feasible. In line with Dorsch (2017) one can propose to be guided by the scientific, academic, or scholarly character of the publishing source and not to consider all those documents which are not published in such media. Pragmatic criteria for a research publication include the existence of an abstract (however, these are not available in every case, e.g., not in editorials); the affiliation of the authors, e.g., universities, research companies, etc. (however, these might be not at hand for independent scholars); the existence of references, if applicable; and the use of a citation style. It seems to be questionable and worthy of discussion to add an institutionalized peer review process by the publishing source as a criterion of a research publication, as the quality of peer review is questionable (Henderson, 2010; Lee et al., 2013). Furthermore, not all research journals perform peer review, and, vice versa, some sources promise peer review and fail to run it satisfactorily. Hence, peer review was seen as a "flawed process" (Smith, 2006, p. 182), but we have to mention that there is no convincing alternative.

One could assume that so-called predatory journals, i.e., online journals without serious peer-review, solely maximizing profit (Cobey et al., 2018), are not scholarly sources. There are different lists of predatory journals (e.g., Beall's List or Cabells' Predatory Reports; Chen, 2019) and also white lists of non-predatory journals, but they do not match (Strinzel et al., 2019). It seems not to be possible to delete articles in predatory journals or proceedings of predatory conferences, as there is no clear cut distinction between predatory sources as pseudo-science and

non-predatory sources as science. It is possible to find even highly cited quality research articles in such sources (Moussa, 2021).

A special problem for scientometrics is retracted papers (Bar-Ilan & Halevi, 2017), i.e., articles which were formally published but later on taken back by the authors or by the publishing source, e.g., due to ethical misconduct, fake data, or false reports. Do retracted papers count as publications for the authors? And do citations to such documents count as equivalent to citations of non-retracted papers? A special problem seems to be the attention paid to retracted papers on social media, as for some retracted articles their attention increases. When papers written by famous authors or published in famous journals “are retracted for misconduct, their altmetric attention score increases, representing the public’s interest in scandals, so to speak” (Shema et al., 2019).

If a researcher or a company asks for patent protection, the document is a patent application, an “A Document,” at first. But if the inventor does not apply for granting the patent or if the patent authority refuses to grant it, is it a research publication in this case? It seems necessary only to consider granted patents, “B Documents,” as research publications (Archambault, 2002). Besides such technical property rights as patents and utility models there are also non-technical property rights such as designs and trademarks. Non-technical property documents, however, as well as utility models in many countries, do not pass through a testing process at the patent office as there is only pure registration without examination. For non-technical intellectual property documents, it seems that there is only a marginal scientific research background, but a strong background in arts (designs) as well as marketing and advertising (trademarks); so it is an open question to include design and trademark documents in scientometric research or not (Cavaller, 2009; Stock, 2001).

In the case of rankings, the research publications used are often limited to single types of publications, for instance, to journal articles (Röbken, 2011). The reasons for this can be limited data availability, e.g., almost only papers in journals are recorded in some sub-databases of WoS. An additional reason can be that certain research rankings only focus on journal articles. As a consequence, only a small fraction of research publications is considered. Besides journal papers, there are monographs, papers in edited books, papers in conference proceedings, patents, and other articles (all other types, such as editorials or reviews) according to Reichmann et al. (2022).

As well, this section ends rather unsatisfactorily, as

we did not find clear criteria to draw the line between research documents as objects of scientometrics and pseudo-science or even non-science publications to be omitted or marked as problematic in scientometric studies. But it should be clear that a set of demarcation criteria would be very helpful for scientometrics. Perhaps this article helps to initialize studies concerning demarcation in scientometrics.

4.2. Data Sources of Research Publications: How to Find Research Publications?

Scientometricians can use personal publication lists of individual researchers, research databases from research institutions (e.g., universities), discipline-specific information services (e.g., Chemical Abstracts Service for chemistry, Inspec for physics, Medline for medicine, or Philosopher’s Index for philosophy), or multi-disciplinary citation databases in order to collect a comprehensive set of publications. Frequently used data sources are the WoS (Birkle et al., 2020), Scopus (Baas et al., 2020), and less frequently applied are Google Scholar (Aguillo, 2012) and (since the end of 2021 inactive) Microsoft Academic (Wang et al., 2020), and the recently established Dimensions (Herzog et al., 2020) as well as OpenAlex, providing “pinnacles” like the ancient Library of Alexandria with access via API (Priem et al., 2022). There are additional sources which are specialized on research data and their citations such as WoS’s Data Citation Index (Robinson-García et al., 2016), DataCite (Ninkov et al., 2021), or Overton to track policy citations (Szomszor & Adie, 2022). Besides this, there exist several repositories, especially utilized as preprint publication data source; well-established ones include arXiv (McKiernan, 2000), medRxiv (Strcic et al., 2022), bioRxiv (Fraser et al., 2020; 2022), Research Square (Riegelman, 2022), and Zenodo (Peters et al., 2017). Additionally, there are discipline-specific preprint services such as Social Science Research Network (SSRN) and SocArXiv for the social sciences.

Concerning the number of research publications, Google Scholar retrieves the most documents, followed by the former Microsoft Academic, Scopus, Dimensions, and the WoS (Martín-Martín et al., 2021). However, all mentioned information services are incomplete when compared with researchers’ personal publication lists (Dorsch et al., 2018; Hilbert et al., 2015). For WoS, Tüür-Fröhlich (2016) could show that there are many misspellings and typos, “mutations” of author names, and data field confusions, all leading to erroneous scientometric results. Concerning the literature of technical intellectual property

rights, i.e., patents and utility models, the information services of Questel and of Derwent are complete for many countries and therefore allow for comprehensive quantitative studies (Stock & Stock, 2006).

In comparison to the other information services, Google Scholar is challenging when advanced searching is required. Since it does not support data downloads, it is difficult to use as a scientometric data source. Additionally, it lacks quality control and clear indexing guidelines (Halevi et al., 2017). For WoS, it is known that it has a bias in disciplines (Mongeon & Paul-Hus, 2016). There is a clear bias in languages in favor of English for WoS (Vera-Baceta et al., 2019), which is also true to a lesser extent for Scopus (Albarillo, 2014). As WoS has many sub-databases and as research institutions commonly subscribe to WoS only in parts (mainly Science Citation Index Expanded [SCIE], Social Sciences Citation Index [SSCI], Arts & Humanities Citation Index [A&HCI], and Emerging Sources Citation Index [ESCI], i.e., sources which only cover journal articles), research disciplines publishing also in proceedings (as does computer science), books (as in many humanities), and patents (as does technology) are systematically discriminated against (Stock, 2021). Besides the WoS Core Collection, we find regional databases for the respective research output from Russia (Russian Science Citation Index; Kassian & Melikhova, 2019), China (Chinese Science Citation Index; Jin & Wang, 1999), Korea (KCI Korean Journal Database; Ko et al., 2011), Arabia (Arabic Citation Index; El-Ouahi, 2022), and South America (SciELO; Packer, 2000), which minimize WoS's language bias. But the scientific institutions and their libraries must subscribe to all sub-databases, which cannot always be realized because of budget reasons. About half of all research evaluations using WoS data did not report the concrete sub-databases (Liu, 2019) and similarly 40% of studies on the h-index withheld information on the analyzed sub-databases (Hu et al., 2020).

Some scientometric studies are grounded on data from an information service (e.g., WoS), which are further processed in an in-house database inside an institution, and which are as a consequence not available for researchers outside the institution (e.g., the study of Bornmann et al., 2015). Hence, an intersubjective examination and replication of such results by independent researchers is not possible, leading Vélchez-Román (2017) to press for general replicability in scientometrics to ensure—paraphrasing Popper (1935)—falsifiability of the results.

There are huge problems with the data sources on all scientometric levels (Rousseau et al., 2018). For Prancutė

(2021, p. 48), there are serious biases and limitations for both WoS and Scopus. At the micro-level (of individual researchers), Dorsch (2017) found massive differences in publication counts between information services, and Stock (2021) was able to present similarly high differences between different library subscriptions of the sub-databases of WoS. Concerning the macro-level (here, universities) Huang et al. (2020) reported on different rankings of the evaluation units depending on the data source. For van Raan (2005, p. 140) many university rankings must be considered as “quasi-evaluations,” which is “absolutely unacceptable.”

Dorsch et al. (2018) also used personal, institutional, or national publication lists as data sources for quantitative science studies. In their case studies, the personal publication lists of authors in scientometrics are much more complete than the research results on Google Scholar, Scopus, and WoS. However, these publication lists are not complete in an ideal way (“truebounded” in the terminology of Dorsch et al. (2018)). In many cases, they are slightly “underbounded,” i.e., there are missing items, and in a few cases, they are “overbounded,” i.e., there are documents not having been formally published or not being research publications. In comparison to the truebounded publication lists, personal publication lists in their study consider about 93% of all publications on average, Google Scholar 88%, Scopus 63%, and WoS Core Collection (only SSCI, SSCI, A&HCI, and ESCI) 49% accordingly. Not every author maintains a current and correct publication list. In such cases, it might be possible to consider institutional repositories (Marsh, 2015) or even data collections on a national level (Sile et al., 2018). Following Dorsch (2017) we must take into account that the entries in the personal, institutional, or national bibliographies will not be in a standardized format, so that further processing is laborious.

However, it can also happen that in the context of rankings, the objectively best, i.e., most complete data source is not used, although it would be accessible, but rather the one the use of which most likely leads to the desired results. For example, instead of publication lists, the WoS could be used to favor researchers who mainly publish in English-language journals. However, if a research evaluation is generally limited to the publication type “papers in journals,” the differences depending on the data source used are much smaller. And if there is another restriction to English-language papers, these differences decrease again (Reichmann & Schlögl, 2021). In general, the use of international databases such as WoS or Scopus

implicitly restricts the research publications retrieved to mainly those being published in English (Schlögl, 2013).

With regard to high-quality professional information services (e.g., WoS and Scopus) one may argue that they only include “quality papers” being published in “quality journals.” However, according to Chavarro et al. (2018), the inclusion of a journal in WoS or Scopus is no sound criterion for the quality of this journal.

Apart from the good situation with patent information, the data sources for other scientometric studies are either relatively complete (as are Google Scholar and some personal, institutional, or national repositories) but poorly suited for quantitative studies; or they are well-fitted for scientometrics (as are WoS, Scopus, or Dimensions), but lack a thorough database. Indeed, many scientometric studies rely on data from WoS and Scopus. For Fröhlich (1999, p. 27), they are “measuring what is easily measurable,” and Mikhailov (2021, p. 4) states that “the problem of an adequate and objective assessment of scientific activity and publication activity is still rather far from being solved.”

4.3. Altmetrics

More than two decades ago, the first social media platforms were launched. Soon after social media also became popular in the world of research, they became used for communication between researchers and between researchers and the general public. This subarea of social media metrics is called “altmetrics,” a term which was originally coined by Priem (2010).

In the following, we will concentrate on basic problems of altmetrics, namely the diversity of data and the diversity of data sources of altmetrics (Holmberg, 2015) and on the reasonableness of altmetrics indicators. Following García-Villar (2021, p. 1), “altmetrics measure the digital attention received by an article using multiple online sources.” With the upcoming of the Internet and digitalization, new ways of informal communication between researchers as well as between researchers and the general public emerged. Besides popular-science descriptions of research results in newspaper or on TV (Albrecht et al., 1994), now, for instance, numbers of views, downloads, discussions in blogs or microblogs, mentions in social networking services, or recommendations can be such new characteristics of research publications (Lin & Fenner, 2013). Download numbers of individual research papers (Schlögl et al., 2014) and contributions in social media such as, e.g., blogs, microblogs, posts, as well as in mass media (e.g., newspapers) especially extended the repertoire of data

sources and research indicators. In their first generation, the analysis of these aspects is summarized as “webometrics” (Almind & Ingwersen, 1997), followed by “altmetrics” (Priem, 2010; 2014) as the second generation of web-based alternative indicators and methods, especially from the social web (Thelwall, 2021).

Altmetrics describes the field of study and the metrics themselves. It is broadly defined as the “study and use of scholarly impact measures based on activity in online tools and environments” (Priem, 2014, p. 266). Since its emergence, the concept itself, the setup of an altmetrics framework, and the methodologies, advantages, challenges, and main issues of altmetrics are discussed. First and foremost, the question is, what do altmetrics really measure? However, this is not the only aspect in question over the last ten years. Priem (2014) already lists a lack of theory, ease of gaming, and bias as concerns. Sugimoto et al. (2017) summarize the conceptualization and classification of altmetrics as well as data collection and methodological limitations in their literature review as challenges. Erdt et al. (2016) show in a publication analysis of altmetrics that aspects like coverage, country biases, data source validity, usage, and limitations of altmetrics were already discussed in research literature in the beginning of altmetrics in 2011. Later normalization issues (2014–2015) and detecting gaming as well as spamming were found as further topics. Coverage of altmetrics and cross-validation studies were the main research topics on altmetrics between 2011 and 2015 (Erdt et al., 2016). Still now, Thelwall (2021, p. 1302) identifies such common issues as “selecting indicators to match goals; aggregating them in a way sensitive to field and publication year differences; largely avoiding them in formal evaluations; understanding that they reflect a biased fraction of the activity of interest; and understanding the type of impact reflected rather than interpreting them at face value.”

Altmetrics are heterogeneous in terms of their characteristics of measurements, the availability of the data, and the ease of data collection (Haustein, 2016; Sugimoto et al., 2017). Depending on the audience, “views of what kind of impact matters and the context in which it should be presented” (Liu & Adie, 2013, p. 31) vary. This overall heterogeneity enables altmetric applications in a variety of different ways and cherry-picking from available data and tools (Liu & Adie, 2013). However, this also makes it difficult to define the impact of the measurements. If one conducts altmetric studies, extensive data cleansing is necessary at any rate (Bar-Ilan, 2019).

If not collected directly from the respective data

source, altmetrics data can be obtained from different data aggregators. Established and frequently used aggregators are Altmetric.com, Plum Analytics, ImpactStory (formerly Total-Impact), and PLOS Article-Level Metrics. Similar to raw data from different sources, these processed data differ (Zahedi & Costas, 2018). In their comparison of the four aggregators, Zahedi and Costas (2018) argue that most data issues are related to methodological decisions in data collection, data aggregation and reporting, updating, and other technical choices. With these inequalities, it is still of utmost importance to understand how differences can cause variations within altmetrics analyses (Zahedi & Costas, 2018). Likewise, the underlying social media platforms affect altmetric indicators in general (Sugimoto et al., 2017). Potential data quality issues can happen at the level of data providers, data aggregators, and users. Furthermore, for providers like Twitter or Facebook, data quality is not a priority as they are not targeted at academia (Haustein, 2016). When altmetrics cover dynamic publications such as, e.g., Wikipedia entries, such texts may be in constant flux. Not uncommon in social media are deleted publications such as, e.g., tweets on Twitter or posts on social networking services. Moreover, hyperlinks included in tweets or posts may not work sometime after their publication. Deleted or modified documents on social media services are a serious problem for the replicability of studies in altmetrics. Nonetheless, Twitter is one of the most popular sources of altmetrics, its data are easily accessible, and the data could be used as early indicators of an article's possible impact (see, e.g., Haustein, 2019).

There are many different sources for altmetrics with different possible measures. One may describe Mendeley readers, Wikipedia mentions, microblogs on Twitter, posts in weblogs, news in Reddit, posts on Facebook, and videos in YouTube, among others. All these measures capture different aspects of communication, and there is not or not always a high correlation between the single measures (Meschede & Siebenlist, 2018). From the perspective of the researcher or person who is using altmetric data, there is the risk of making application errors. Following Thelwall (2021), using a “basket of indicators,” or “a single hybrid altmetrics,” without a theoretical or problem-based reason instead of using indicators tailored to the research framework and goals is not uncommon. Likewise, the usage of sparse indicators for too small datasets is not recommended. Just as with conventional scientometrics, ignoring field differences or publication year differences would affect the analysis. For example, the comparison of two fields where one field is much more engaged and

discussed within social media would require field normalized indicators. Calculating arithmetic means or using Pearson correlations to validate altmetrics indicators without log-transformation or normalization is also not recommended, as altmetrics, like webometrics and citation counts, are highly skewed. It is furthermore suggested not to over-interpret correlation values. In general, altmetric indicators are also biased and only “reflect a very small fraction of impacts of the type that they are relevant to” (Thelwall, 2021, p. 1308). For more general indicators like Twitter citations, it is important to attribute meaning with, e.g., content analysis of citation sources, surveys, or interviews instead of “equating the type of impact that they might reflect with the type of impact that they do reflect” (Thelwall, 2021, p. 1308). Another aspect is to analyze indicators separately instead of using hybrid indicators (e.g., Altmetric Attention Score) that aggregate several altmetrics into a single score, as the latter have different interpretations. Altmetrics data is in constant flux, so not reporting on the date of data collection and how the data were collected would further add against interpretability. It is also important to conduct an effective literature review to keep updated about issues and directions in altmetrics. There is always the possibility that a social media service applied for altmetrics such as, for instance, Twitter, will be shut down or no longer used by researchers. In terms of formal research evaluations, altmetric indicators should be avoided (Thelwall, 2021).

Altmetric measures provide both further characteristics of formal research publications besides citations, for instance, number of views, downloads, or recommendations, and further research-related informal publications in social media as well as in newspapers, TV, or radio. Overall, altmetrics share many characteristics, and thus problems, with citations so that methodological solutions already exist and can be imported (Sugimoto et al., 2017). Altmetrics add to the bouquet of a research publication and can contribute to gathering more information on different societal impacts of research (Thelwall, 2021). Herb (2016) objects that altmetrics refrain from every serious theoretical foundation; further—for Herb (2016, p. 406)—it is unclear which properties of which objects with which operationalizations altmetrics really measure. There is a huge problem of altmetrics: Altmetric measures capture different aspects of the informal internal as well as external research communications—the problem is how to merge all these aspects as there is no standard (Holmberg, 2015). Perhaps there should be no merging of different data from different data sources at all. While such single

measures may be indeed okay, there are serious problems with indicators derived from those measures. Especially, aggregated indicators are more or less senseless, as we do not know what they exactly tell us.

5. WHAT IS ONE RESEARCH PUBLICATION?

In this section, we discuss the counting units of scientometrics. What causes us to count “1” for a research publication, reference, citation, or topic? Counting problems can arise from multi-author publications as well as from weightings for publications. We also discuss the units of references and citations and of topics. Finally, we deal with counting problems on the meso- and macro-level of scientometrics.

5.1. Multi-author Publications and Their Counting

If a research publication has one author, then this publication counts as “1” for the researcher. With multi-author publications, counting problems start (Gauffriau, 2017), as counting methods are decisive for all measures of authors’ productivity and for rankings which are based upon publication and citation studies (Gauffriau & Larsen, 2005). What is the share of each author? Egghe and Rousseau (1990, pp. 275-276) differentiate between straight counting (considering only first authors and ignoring all co-authors), unit counting or whole counting (for all co-authors the publication counts “1”), and fractional counting ($1/n$ in the context of n co-authors). The formula $2/(n+1)$ has advantages in terms of a balanced consideration of co-authorship in relation to single authorship (Schneijderberg, 2018, p. 343). It takes into account the fact that co-authorship requires additional work due to the necessary coordination. However, the fact that the sum of the author values is greater than one in this variant is a decisive flaw.

There are further weighting methods considering different values for different kinds of contribution following the Contributor Roles Taxonomy (Rahman et al., 2017). Other approaches consider the past roles of the co-authors such as, for instance, senior versus junior authors (Boxenbaum et al., 1987), the co-authors’ h-indices (Egghe et al., 2013), or the co-authors’ numbers of publications and received citations (Shen & Barabasi, 2014). All in all, Gauffriau (2017; 2021) found 32 different counting methods for an author’s share of a multi-authored research publication. Among them are problematic practices as well, like the authors’ h-index.

Straight counting of only the first author is unfair towards all authors who are not the first author. If we

want to weight contributorship fairly, we need weighting schemes for all kinds of contributions and aggregation rules for authors with different contribution types. However, a practical problem arrives as no information service provides contributor data for now. The consideration of the authors’ histories with, e.g., the amount of citations, their position in the institution, or their h-indices, is in particular unfair towards younger researchers.

Anyway, unit counting and fractional counting are fairer methods, which do not discriminate against any of the co-authors. However, when we aggregate data from the individual researcher level to higher levels, unit counting may lead to wrong results. If we assume that two authors from the same institution co-publish an article and count “1” for each author, this article would count “2” for the institution. As co-authorship of researchers from the same institution, city, country, etc. is a matter of fact, the adding of the single values would result in too high values at the aggregate level. This mistake can be prevented by relating whole counting directly to the aggregating level (Reichmann & Schlögl, 2021).

Even when applying fractional counting, it must be considered that proportional authorship according to the formula $1/n$ does not reflect the “true” contribution of every author; it is a kind of leveling down of each co-author. An information science article with, for instance, two authors counts $1/2$ for each; a high-energy physics article with 250 authors (which is thoroughly common there) counts $1/250$ for each. This makes every comparison between disciplines with different publication cultures nearly impossible. A solution can be a field-specific normalization of co-author values (Waltman & van Eck, 2019). However, such a normalization is a difficult task, as there are no realistic normalization factors for single disciplines, as one must consider not only the research area, but also the kind of papers, such as, e.g., a research paper versus review article. Furthermore, the underlying distributions are usually skewed, leading to problems of the calculation of arithmetic means or median values (Kostoff & Martinez, 2005).

Many research journals ask for research data if the article reports on empirical findings. If the research data, say, the applied questionnaire or sheets of numerical data, is published in a digital appendix of the paper, the unit of article and data counts “1.” But what should be counted if the research data is published in another source such as, e.g., on Zenodo? Should scientometricians ignore the publication of the research data when counting an author’s publications? A possible solution would be a separate

counting of research articles and research data. The same course of action can also be applied for the calculation of the impact. Accordingly, the citations and mentions in social media on those research data could also be observed apart (Peters et al., 2016). As information services for research data do not check the data's quality, they must be regarded as informally published. However, since they are connected with their formally published research paper, they are in between formally and informally published documents.

It is clear that different counting procedures “can have a dramatic effect on the results” (Egghe & Rousseau, 1990, p. 274). In scientometrics, unit counting and fractional counting are often applied; however, both approaches lead to different values (for the country-level see Aksnes et al., 2012; Gauffriau et al., 2007; Huang et al., 2011). Of course, rankings derived from different counting approaches cannot be compared (Gauffriau & Larsen, 2005). Additionally, there are scientometric studies which do not even give detailed information on the applied counting approach (Gauffriau et al., 2008). For Osório (2018, p. 2161) it is impossible to name a general counting method able “to satisfy no advantageous merging and no advantageous splitting simultaneously.” Consequently, the problem of the unit of publication for multi-authored research publications remains far away from a satisfying solution.

5.2. Weighting Publications

A book of, say, 800 pages by two authors has with fractional counting a value of 0.5 for each of them; a two-page review of this book by one author counts 1.0. Is this fair? A citation of this monograph in a long article on this work in, for instance, *Science*, counts 1.0; and a citation without substantial text in a rarely read regional journal counts 1.0, too. Is this fair? If the answers for those questions are “no,” scientometric studies have to work with weighting factors for research publications and citations.

How can we weight different publication types? In a rather arbitrary way, some authors have introduced weighting factors, for instance, 50 for a monograph, 10 for an article, and 1 for a book review (e.g., Finkenstaedt, 1986). These weighting factors are often based on the average size, which is why monographs usually receive the highest number of points. Papers in journals sometimes receive a higher number of points than those in edited works (Reichmann et al., 2022). However, such document-specific weighting approaches were not able to find broad application, because all introduced weighting factors are more or less employed arbitrarily.

Should a book be weighted with its length (Verleysen & Engels, 2018), with the number of published printed copies (Stock, 2001, p. 21), or with download numbers? Is a book with different chapters one publication, or is every chapter also one publication each? If an information service such as, for instance, WoS in its Book Citation Index, indexes chapter-wise, every chapter counts as a publication. The same procedure is usually applied in the case of edited books. The question remains whether the mere editing should also count as a publication for the editor.

If the number of published pages is used for taking into account the size of a publication, there can be strong distortions due to different formats, font sizes, and layouts. It would be better to count the size in words or characters, but such data is only available for digital publications. The size of publications can also be used as a filter criterion, for example, by only considering publications with a certain minimum size in a research evaluation (Reichmann et al., 2022). There are no international standards for the optimal length of a publication, which will certainly depend heavily on the type and medium of a publication, but scientific contributions, especially those in journals and edited works, should have a certain minimum length in order to be regarded as full-fledged. However, if there were, for example, a minimum size of five printed pages, many articles in highly respected journals such as *Nature* or *Science* would no longer be considered.

Since the papers by Seglen (1992; 1997) it is known that an article's weighting with the impact factor of the publishing journal is misleading, since impact factor is a time-dependent measure for a journal and not for individual articles. Furthermore, if we sort articles in a journal by their citation counts, we can find a highly left-skewed distribution with few heavily cited papers and a long tail of articles being much lesser or even not cited at all. As the impact factor is a quotient of the number of citations to a journal in a given year Y and the number of source papers in the same journal in the years Y-1 and Y-2 (Garfield, 1979), journals which publish primarily up-to-date results (i.e., from the last two years), and therefore are cited soon afterwards, are heavily favored in contrast to journals with longer dissemination periods such as, for instance, in disciplines like history or classical philology. Quantitative descriptions of journals such as the impact factor can lead to a kind of halo effect, i.e., the link of a current evaluation to previous judgments, which is a cognitive bias (Migheli & Ramello, 2021). The impact factor depends on only a few highly cited articles which are generalized to all papers, so the highly cited papers generate the halo effect for

the less-cited ones. Accordingly, the meaning of an impact factor is “systematically biased” (Migheli & Ramello, 2021, p. 1942). The application of the journal impact factor on individual papers was called a “mortal sin” (van Noorden, 2010, p. 865, citing a statement by van Raan); however—and here the application of pseudo-scientometric results becomes problematic—one can find appointment committees in universities working with such misleading indicators (Martyn, 2005).

Similarly, it is not without problems to use discipline-related journal rankings like, for example, JOURQUAL 3 for the field of business administration (Schrader & Hennig-Thurau, 2009) for the weighting of articles. However, this approach also ignores the fact that it is hardly possible to infer the quality of a single paper from the quality of the journal in which it is published.

As already mentioned, granted patents can also be counted as research and development publications. A weighting factor for patents would be the extension of the respective patent family, i.e., the number of countries for which the invention is granted (Neuhäusler & Frietsch, 2013). Additionally, one can weigh the patents by the economic power of the granting country (Faust, 1992), so a patent granted in the United States weighs more than one granted in, for instance, Romania or Austria.

To sum up, we did not find hints for weighting factors for research publications, which are generally acknowledged by the scientometrics community as all factors represent more or less arbitrary decisions.

5.3. The Unit of References and Citations and Weighting Citations

A similar problem as determining the unit and weighting of publications comes with the unit and weighting of references and citations. What is one reference and the respective citation? It is common practice in scientometric studies and in citation databases to count every different reference in a publication as a citation of the mentioned publication, so we count the fact of information dissemination from one publication to another. But we do not count the number of references to a cited publication in the text or in footnotes with “loc. cit.,” “ibid.,” etc. (Stock, 1985). Let us assume that in a hypothetical paper there are two references XYZ, 2021 and ABC, 2020. In the paper’s body, ABC, 2020 is mentioned ten times, but XYZ, 2021 only once in a short sentence. Intuitively, we would count ABC, 2020 ten times and XYZ, 2021 once. However, every citation database counts only “1.” For such in-text citations, Rauter (2005) and later Zhao et al. (2017) introduced

frequency-weighted citation analysis, counting not only “10” in our hypothetical case for ABC, but also the sentiment of the citation (supportive, negative, etc.). Concerning negative citations, Farid (2021, p. 50) states that they have to be excluded for further calculations. However, there is a large practical problem: how can we determine a citation’s sentiment? Should every author mark every citation in their article as “positive,” “negative,” or “neutral,” as Rauter (2005) suggests? Or should scientometrics work with automatic sentiment analysis in the sentences in the article’s body, where the reference is found (Aljuaid et al., 2021)? Additionally, one could locate the citation in the article, for instance, in the introduction, methods, results, or discussion section (Ding et al., 2014).

How can we weight citations? Self-citations play a special role in citation analyses, as researchers can inflate their citation rate through massive self-citations. It is not useful to count self-citations with “0,” as it may be very informative that an author uses former insights in a later publication. Additionally, self-citations may signal continuity and competence in a particular research field. Schubert et al. (2006, p. 505) suggest weights for author self-citations following the kind of self-citation. Accordingly, a self-citation link between a citing and a cited work is stronger if a single-authored paper is cited in a later single-authored paper by the same author than between two multi-authored papers where the citation is made by only one joint co-author (Schubert et al., 2006, p. 505).

Citations of co-authored publications can be counted “1” for each co-author (whole citation counting) or—using analog to fractional counting of publications— $1/n$ for every co-author (fractional citation counting) (Leydesdorff & Shin, 2011).

As not all citing sources are equally important, one can weight a citation by the importance of the citing articles or by the importance of the citing source. It does not seem to be a good idea to work with the journal impact factor of a citing article as this is the same dead end as in the case of the weighting of an individual article with the journal’s impact factor. But it is possible to weight a citation with the importance of the citing articles. Here, network analysis comes into play. Many scientometric studies work with PageRank (Brin & Page, 1998) or similar eigenvector-based measures of the citing articles (e.g., Del Corso & Romani, 2009; Ma et al., 2008), which are, however, very time-dependent.

In a review on citation weighting solutions, Cai et al. (2019, p. 474) state that “there isn’t a perfect weighting approach since it’s hard to prove that one indicator out-

performs another.” Additionally, the weighting values of citations (and, not to forget, the absolute numbers of citations) change starkly in the course of time. There is always a cold-start problem in the citation history of a research publication, as in the beginning there is no citation: The publication has to be read, and the citing publication has to be written and formally published—a process which takes months or even some years.

5.4. The Unit of Topics

How can scientometricians determine what a topic is? What is the unit of a topic? And do weighted topics exist? As there are different methods of the representation of a research publication's aboutness, there are also different answers to our questions. We will discuss scientometric topic analysis based on title indexing, text statistics, citation indexing, author keywords, index terms (nomenclatures, thesauri, and classifications), text-word method, and content analysis. The methods of title indexing, text statistics, citation indexing, and author keywords are able to capture the author aboutness, while index terms, text-word method, and content analysis describe the indexer aboutness of a publication.

Topic identification, which is based upon text statistics (including the statistics of title words), works with identified single terms and phrases or alternatively with *n*-grams in the text (Leydesdorff & Welbers, 2011). To create relationships between terms, phrases, or *n*-grams, in many cases co-word analysis and co-word mapping are applied. One may weigh the importance of terms, etc. following the frequency of occurrences in both the single publications and in a set of documents (e.g., a hit list of an information service). Further steps of processing may be calculations of term similarity or a factor analysis of the terms. The main problems of text statistics are different publication languages, different language usages of the authors, linguistic challenges such as, e.g., lemmatization or stemming, phrase detection, name recognition, the selection of “important” topics, the choice of the data processing methods such as, for instance, similarity calculation, the creation of topic clusters, and, finally, the presentation of the clusters in an expressive image. An additional problem for title indexing is that there are only few terms for processing. However, an advantage of text statistics for the task of topic identification is that all steps can be automatized.

For Garfield (1979, p. 3), citations represent concepts and so topics. However, with new citations of the same publication the represented concept and topic may change,

as the citing papers come from different subject areas. In addition to the analysis of direct citations (Garfield et al., 2006), scientometricians can apply co-citation and bibliographic coupling for topic identification (Kleminski et al., 2022). Honka et al. (2015) describe a combination of citation indexing and title indexing. In their topical impact analysis, they analyze the title terms of the citing publications of all the literature of a research institution. Sjögarde and Didegah (2022) found associations between citations and topics insofar as publications covering fast-growing topics are more cited compared to publications in slow-growing or declining topics. However, we have to recall that if we apply citation analysis for topic analysis, we inherit all discussed problems of citation analysis.

Keywords are often given by the authors themselves, but information services like WoS also generate so-called “KeyWords Plus” automatically, which are single words or phrases that frequently appear in the titles of an article's references besides the author keywords (Haustein, 2012, pp. 90-91). Both types of keywords include hints to research topics (Li, 2018) and complement title indexing. But author keywords depend heavily on the author's language usage and WoS's KeyWords Plus depends on the author's selection of cited papers (if there are references at all).

Index terms are found in bibliographic records of library holdings or discipline-specific information services such as Medline or PubMed (medicine) or Derwent (patents). Nomenclatures (Stock & Stock, 2013, pp. 635-646) or thesauri (Stock & Stock, 2013, pp. 675-696) such as the MeSH reflect the topics of the sources of the respective service. For Shu et al. (2021, p. 448), MeSH or the controlled vocabulary of similar nomenclatures and thesauri are well suitable for mapping science using subject headings. Concerning classification systems (Stock & Stock, 2013, pp. 647-674), the International Patent Classification (IPC) is often used for topic analyses. Scientometric—or better, patentometric—studies apply both the notations of the IPC and the co-occurrences of those notations (da Silveira Bueno et al., 2018). WoS offers with its Subject Categories a very rough classification system for the topics of covered journals (Haustein, 2012, pp. 94-101), which is indeed applied in scientometric topic analyses, mainly for high-level aggregates such as cities and regions (Altvater-Mackensen et al., 2005). As journals in WoS may be assigned to different classes, this has also a certain effect on topic identification (Jesenko & Schlögl, 2021).

Content analysis works with the classification method, for which the (as a rule: few) classes are built with respect

to the set of publications which should be analyzed. The class construction is both inductive (derived from the publications to be analyzed) and deductive (derived from other research literature). The construction of the classes and the sorting of the publications in the classes is a complex procedure, which is also sometimes used in scientometrics (Järvelin & Vakkari, 2022).

While nomenclatures, thesauri, and classification systems work with a standardized controlled terminology, the text-word method (Stock & Stock, 2013, pp. 735-743) only allows words from the concrete publication as index terms, making this indexing method as low-interpretative as possible. For Henrichs (1994; see also Stock, 1984), the text-word method is ideally suited for contextual topic analyses and for the history of ideas.

For topic analyses, there are ways to weight the topics in a research paper. If one applies MeSH, one can work with a binary weight, i.e., major and non-major topics. The text-word method as a form of syntactical indexing weights every index term in the interval between greater 0 and 100 while calculating the importance of the term in the text (Stock & Stock, 2013, pp. 766-767). If an information service works with coordinating indexing, every index term forms a topic unit and weights “1.”

Indeed, there are many approaches to describe topics in scientometrics. But there is no widely accepted “golden path” to scientometric topic analysis. As the concept of “aboutness” itself is rather blurred, so are topic analyses. However, if we work in the framework of one method such as with classification systems or with thesauri, and with exactly one tool such as, e.g., IPC or MeSH, there should be reliable results for the analyzed topic domain (in our examples, patent literature or medicine). Cross-frontier comparisons between domains, methods, and concrete tools are not possible unless there are semantic crosswalks between different vocabularies. Sometimes mapping is applied, where the concepts of different tools, for instance of different thesauri, are set in relation (Dorerr, 2001). It is also possible to create a crosswalk between different methods, e.g., between the Standard Thesaurus Economy and the classification system of Nomenclature statistique des activités économiques dans la Communauté Européenne (NACE) (Stock & Stock, 2013, p. 724). Nevertheless, such crosswalks are scarce and their applications in scientometrics are nearly unknown outside of patent analyses.

All in all, the scientometric description, analysis, and evaluation of topics of research publications is still a great challenge, as there are no broadly accepted methods and

also no accepted indicators.

5.5. Counting Problems on the Meso- and Macro-levels

Scientometrics on the meso- and macro-levels can apply size-dependent indicators (irrespective of the size of a research unit) or size-independent indicators (considering the size of a unit of research) (Waltman et al., 2016). It seems to be unfair to compare different research institutions (scientometric meso-level) or even countries (macro-level) through their absolute numbers of research publications and citations, since this favors large institutions or countries. Nevertheless, this often happens, for example, in the case of university rankings (Olçay & Bulu, 2017). For all studies on the meso- and macro-level which use whole counting it must be taken into account that co-authors from the same institution or the same country are not counted multiple times. A possible solution might be to relate the fractional counts to the total number of relevant researchers (as did, e.g., Abramo & D’Angelo, 2014 and Akbash et al., 2021) or—better—to the relevant full-time equivalent numbers (Kutlača et al., 2015, p. 250; Stock et al., 2023) for the whole analysis period (Docampo & Bessoule, 2019).

In contrast to evaluations of single researchers, scientometric studies on the meso- and macro-level have, first of all, the challenge of capturing the exact number of staff members in the institution or country in the considered time frame (Toutkoushian et al., 2003) as the “underlying production capability” (Thelwall & Fairclough, 2017, p. 1142). In this context, it must also be clarified whose research performance should be relevant for the research evaluation. The answer to this question seems simple at first glance, namely that of the evaluated units, i.e., in the case of a research evaluation at the institute level, that of the compared institutes (Rousseau et al., 2018). If the research performance is to be recorded as comprehensively as possible, all researchers, i.e., all scientific employees, must be taken into account. However, it must be considered that the research performance varies strongly between research staff. Pre-doctoral researchers, for instance, sometimes have hardly any publications because they often need several years to familiarize themselves with research. Therefore, it could make sense only to consider professors and post-doctoral researchers as representatives of a research institution. There is an additional problem with the acquisition of the researchers’ personal data in particular for longer time periods (Reichmann & Schlögl, 2022), as such data are not always and everywhere

freely accessible, and the access may be restricted due to national data protection acts.

6. KEY FINDINGS AND OUTLOOK

The main contribution of this article is a compilation of problems in scientometrics and research evaluation when it comes to the exact determination of the counting units of these knowledge fields. There is only a little completely new in the paper as most issues can be found in the literature, but the concise synopsis reveals several serious problems.

6.1. Main Results: A Dozen Problem Areas

The key findings of our paper can be subsumed as follows: Problems with scientometric analyses already start with basic assumptions that a publication is (1) formally or informally published, that it (2) contains research content, and that (3) in the publication's byline, the names of the authors are stated. The concepts of (4) references and citations and of (5) topics and their descriptions are also not free of fundamental problems. Coming to the aspect of "research" in research publications, we found three problem areas: the criterion of demarcation (extension of 2), (6) the data sources of formally published publications and their problems, and (7) the data sources of informally published publications and their problems. Finally, there are challenges concerning the counting unit of research publications, references or citations, and topics. What exactly are 1 publication, 1 reference, 1 citation, and 1 topic? Identified problem areas here are (8) multi-author publications, (9) weighted publications, (10) the unit of references and citations, (11) the unit of topics, and (12) additional counting problems on the meso- and macro-levels. These problem areas are numbered according to their first appearance in this paper. This sub-section also provides the answers to our three research questions.

6.1.1. *What are a publication, a reference or citation, and a topic?*

(1) Formal publishing happens through classical communication channels including journals, proceedings, and patents; informal publishing uses newspapers as well as other mass media sources and sources of the social web such as Mendeley or Twitter. The classical formal channels have been studied by scientometrics for decades, while the observation of the digital informal channels was added as "altmetrics" with the upcoming of the social web with its user-created content. Articles on preprint servers and

published research data cannot be regarded as formally published as there is no gatekeeping process, but they are also not informally published as they are connected with their respective formal publications, if they exist at all. So they are "in between" formal and informal publishing.

Publications are—in the sense of Briet (1951/2006) and later Buckland (1997)—documents. However, documents are not only papers, but sometimes also artifacts as buildings as documents of engineering or software programs and information systems as documents of computer science, which are not captured by scientometric methods at all. Published non-artifact documents may be analyzed as works, expressions, manifestations, or items, but scientometrics makes these important distinctions only rarely. Since there are different field-specific cultures concerning publications, citations, and mentions in newspapers and social media, scientometric and altmetric analyses have to consider these differences, e.g., by normalization; however, we miss broadly accepted methods.

(2) Research content calls for the criterion of demarcation between research and non-research (see below!).

(3) Authorship is partly indeterminable as there are ghost writers and also honorary authors. As not all authors contribute equally to a publication, it could be wiser to work with contributorship instead of authorship. But also, in this case it is not determinable whether the declarations of contributorship are correct. Moreover, there is a practical problem as no bibliographic information service contains such data.

(4) References are bibliographic descriptions of documents used in publications and so an expression of intertextuality. For an individual publication, intertextuality is an information flow from an older publication to the actual one as shown by the reference; in the opposite direction intertextuality is considered as a citation and is regarded as reputation for the cited publication. The usefulness of reference and citation analyses clearly depends on the veracity of the references given. And this veracity is not or not always prevalent, as authors cite publications which were not used in the actual paper, and they do not cite all influences. Furthermore, contributions by practitioners in research journals are less regarded in citation studies since they do not cite to the same degree as academics.

(5) Topics are descriptions of the publication's aboutness. The author aboutness can be analyzed by means of text-statistical methods with all the problems of author-specific language use. The indexer aboutness can be identified by applying the index terms from discipline-specific databases. Here, the method (e.g., classification or

thesaurus), the vocabulary of the concrete tool (e.g., IPC or MeSH), the tool's indexing rules, and the work of an indexer play crucial roles. All the steps in the process of the description of topics are prone to failure.

6.1.2. *What is a research publication?*

(2 in detail) In the philosophy of science, the problem of defining "research" in contrast to non-research or pseudo-research is discussed as the criterion of demarcation, but this discussion did not find an accepted solution. In scientometrics, we were not able to find serious discussions on the demarcation criterion at all. A criterion for scientometrics could be to bank on the claimed or supposed research character of the publishing sources. For instance, scientometricians could rely on journals or proceedings meeting basic requirements to ensure quality standards like, for instance, peer review. But peer review may work badly, and there is the additional issue of predatory journals and conferences. It is also unclear how to handle papers on preprint servers, which are not published on a classical channel, and how to count published research data.

(6) There are many data sources which can be used for scientometric studies including multi-disciplinary information services such as, e.g., WoS, Scopus, Dimensions, or Google Scholar, and discipline-specific databases as Medline for medicine. All named sources have problems concerning completeness. WoS consists of diverse sub-databases, which are differently subscribed by research libraries. Some sources do not satisfactorily support scientometric analyses (such as Google Scholar). Alternatively or additionally to information services, scientometrics can apply the data of personal, institutional, or national publication lists. However, these data sources are also not fully complete and the data are not in a standardized format.

(7) Some of the social media sources, for instance, tweets in microblogging services, mentions in Wikipedia, references in Mendeley, posts in other sources with user-created content, download numbers, and—not to forget—articles in newspapers or mentions in TV shows—may give supplementary scientometrics information. However, documents on social media may be deleted later or included links may lead to nowhere, so replications of studies in altmetrics are partly not possible. One may interpret measures of altmetrics as indicators of research-internal and research-external interestingness, but a theoretical justification is still pending.

There are some online sources which are easily accessible and the data of which are well processable, e.g.,

WoS, Scopus, and Dimensions for formally published documents and Twitter for informally published documents, which are frequently used in scientometrics and altmetrics. None the less, scientometricians should always use those data sources which are best suited for their scientometric problems to be investigated (e.g., researchers' personal publication lists) instead of easy-to-edit sources.

6.1.3. *What are one research publication, one reference or one citation, and one topic?*

(8) Scientometricians may apply different counting methods for determining the contribution of authors to an individual article. Often used are unit counting (all authors count 1) and fractional counting (e.g., counting $1/n$ for every co-author given n names in the byline). Different counting methods can have a huge effect on the findings.

(9) It is evident that a book with twelve chapters cannot be counted as 1, whereas the twelve chapters, published as articles, would count for 12. However, there are no valid and reliable weighting methods so far. Many approaches are arbitrary (as the weighting of publications by document type) or simply wrong (e.g., applying the journal impact factor to individual papers). For problems no. 8 and 9, we were not able to find an accepted gold standard in scientometrics.

In scientometrics, (10) a reference is considered as an information flow from an earlier document to the citing one, while not considering the extent of the information flows. If an article references another paper, for instance, 20 times and another one only once in a short note, both count—against evidence—1. Citations of multi-authored publications are normally counted 1 for every cited author; but fractional citation counting (e.g., $1/n$ for every co-author) would represent a much fairer solution. We also found approaches for weighted citations as, e.g., using the PageRanks of the citing articles, but no perfect solution.

In order to capture (11) research topics for quantitative scientometric descriptions and analyses, there are indeed many paths including text-statistical approaches and the application of index terms from bibliographical information services. We found options for binary topic weighting (for instance, major versus non-major topics) and interval-scaled weighting (by applying a database working with syntactical indexing). However, we failed to identify a widely accepted method for topic description and for topic weighting.

On the meso- and macro-levels of scientometrics and especially of research evaluation, special problems arise

(12): Who is a member of an institution? Should research evaluation consider the institution's size? If yes, do we count faculty or all staff? Should full-time equivalents be used instead of head counts? Is it possible to get these numbers at all due to data protection regulations?

6.2. Counting Problems of Scientometrics in an Overview

In Fig. 1, we summarized important problems regarding the empirical basis of scientometrics and research evaluation. Additionally, we categorized the problems into three areas. Firstly, there are problems of the research system such as, for instance, authorship, redundant publications, and veracity of references. These problems are caused by the research system, and scientometrics is not able to solve them alone. The second problem category includes the areas of data collection and determination

of the counting units of the publications, references, citations, and topics such as the handling of multi-authored publications and their citations, normalization, the weighting of publications and citations, and the methods for capturing topics. We miss clear definitions of formally published documents, informal published documents, and documents in between such as preprints and research data. The last category includes problems with scholarly indicators. As there is a gap between measures and indicators, the interpretation of the indicators is essential. Measures, e.g., counts of citations, are safe terrain, apart from the problems listed above and apart from measurement errors. Indicators refer to not directly observable constructs. For instance, counts of citations are regarded as indicators of impact in the research community, which is a construct. They allow for interpretations and are therefore open for uncertainties. The inference from an

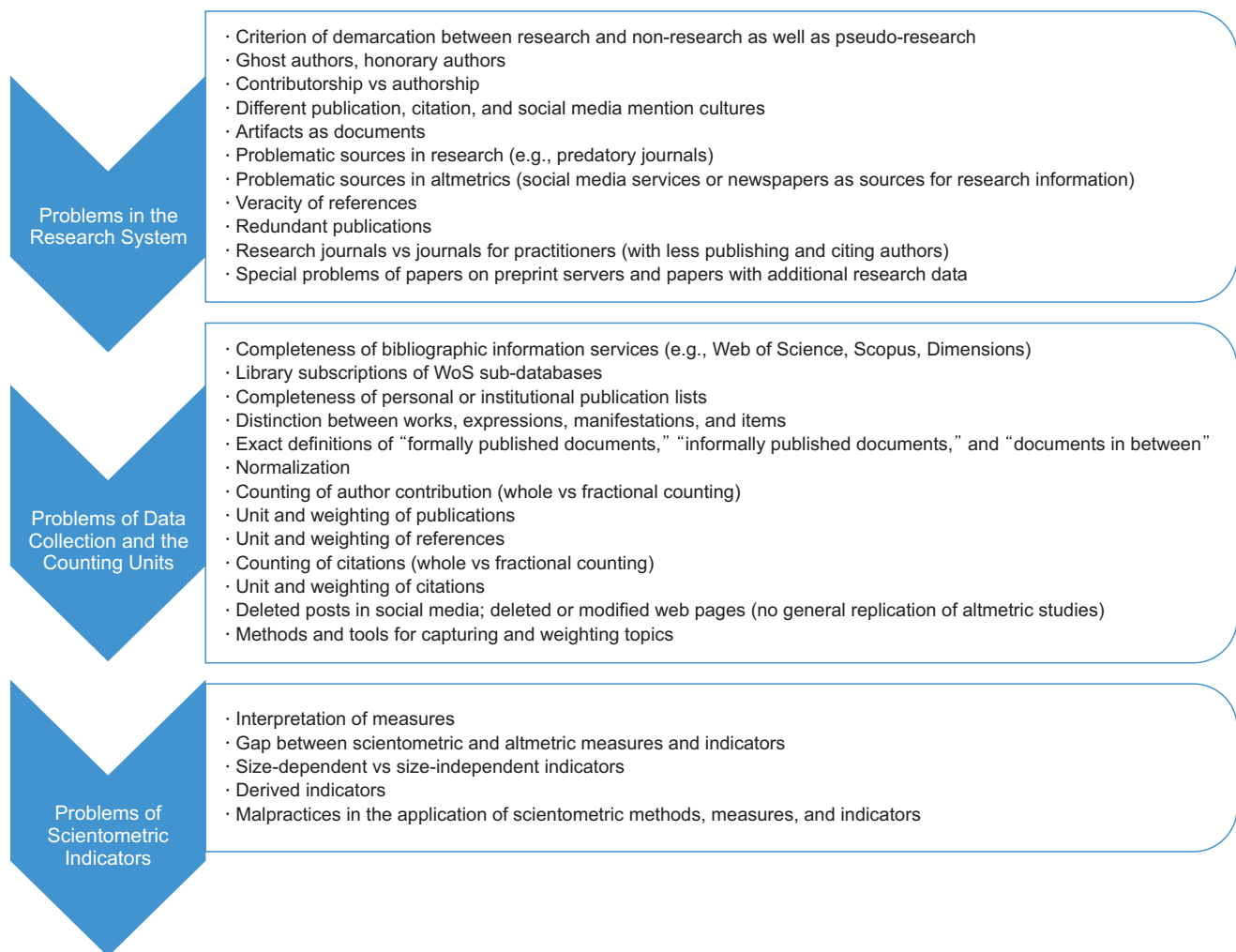


Fig. 1. Problem areas with regard to the empirical basis of scientometrics and research evaluation.

indicator of a thing to the thing itself is called “reification,” which should be avoided (Stock, 1994, p. 24; Weingart & Winterhager, 1984, p. 26). Therefore, “indicators built from patents and articles cannot be perfect” (Narin et al., 1994, p. 69). Besides the interpretation of measures and the mentioned gap between scientometric and altmetric measures and their indicators, there are the problems concerning size-dependent and size-independent indicators and the problem areas of derived indicators.

Quantitative scientometric studies are therefore by no means unproblematic. Though problems in the research system are outside of scientometrics, these issues are reflected in scientometric studies. The problems increase when it comes to the correct data collection including the counting of research publications, citations, and topics. Finally, there are problems with scientometric indicators and their interpretation. Indeed, it seems dangerous to rely only on scientometric indicators for the description, analysis, evaluation, and policy of research. Additionally, all mentioned problems may lead to malpractices especially in research evaluation, when evaluators are driven by research-external interests and prefer measures or indicators which primarily meet their interests.

6.3. Outlook

We concentrated on the counting unit of scientometrics and research evaluation, namely the research publication, its measures, and its indicators. When analyzing the foundations of derived indicators, we would probably find further problems. What are the strengths or weaknesses of, for instance, the impact factor, the h-index, or the altmetrics score? What do they really measure? Are researchers aware of scientometric and altmetric indicators (Keng et al., 2022)? How do scientists appraise the indicators’ importance for themselves and for their knowledge field (Kamrani et al., 2020; 2021)? Do the scientists have a sufficient level of metrics literacies to be able to understand all measures and indicators and their problems (Dorsch et al., 2021)? Of course, there are several articles on these topics. However, there is room left for further reviews and for detailed criticism about the basics of scientometrics.

Every scientometric study should take into account the problem areas mentioned in this article as much as possible. One should be aware, however, that compromises always have to be made in such analyses. Nevertheless, quantitative scientometric studies usually have a certain benefit and added value. However, it is absolutely necessary to point out the limitations of such a study and, ideally, to rethink the scientometric results by conducting an

additional qualitative analysis. We have learned from Leydesdorff (1995) that the science of science is broader than scientometrics as it also comprises the sociology of knowledge and theories of information and communication. For Leydesdorff (1995, p. 10), there is room “for the use of qualitative data and dynamic analysis” in addition to purely quantitative scientometrics. Therefore, scientometricians should overcome the divide between quantitative and qualitative studies in the science of science (Cambrosio et al., 2020; Kang & Evans, 2020; Leydesdorff et al., 2020). Qualitative methods in science studies cover, for instance, interviews with researchers, content analysis of research publications, ethnographic field research (e.g., observations of researchers at work), surveying researchers (also with open questions), and grounded theory (in order to arrive at theories on research communication) (see from another knowledge domain: Gremm et al., 2018, pp. 75-78).

Furthermore, the adoption of theoretical approaches from communications science (e.g., the Uses and Gratifications Theory or the Lasswell Formula of Communication; see Zimmer et al., 2018) is helpful for the analysis of scholarly communication in scientometrics, too. With the help of such approaches, we may analyze, for instance, in the context of the Lasswell Formula (Lasswell, 1948), the way from research papers’ authors (who?) to the intended target group (to whom?), considering also the content (says what?), the paths of transmission (in which channels?), and the readers’ reactions, e.g., applying or citing the paper (with what effect?). Or we may analyze the research communication from the opposite direction: Why do people read, apply, and cite a research paper? For the Uses and Gratifications Theory by Katz et al. (1973-1974), there are seven aspects in the audience’s research usage: (1) the social and psychological origins of (2) the audience members’ needs, which generate (3) expectations of (4) the research literature or other sources, which lead to (5) different patterns of research exposure, resulting in (6) need gratifications and (7) other consequences (including unintended ones). In this manner, scientometrics will not only count and interpret publications, topics, and citations, but the entire process of scholarly communication.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

REFERENCES

- Abramo, G., & D'Angelo, C. A. (2014). How do you define and measure research productivity? *Scientometrics*, 101(2), 1129-1144. <https://doi.org/10.1007/s11192-014-1269-8>
- Aguillo, I. F. (2012). Is Google Scholar useful for bibliometrics? A webometric analysis. *Scientometrics*, 91(2), 343-351. <https://doi.org/10.1007/s11192-011-0582-8>
- Akbash, K. S., Pasichnyk, N. O., & Rizhniak, R. Y. (2021). Analysis of key factors of influence on scientometric indicators of higher educational institutions of Ukraine. *International Journal of Educational Development*, 81, 102330. <https://doi.org/10.1016/j.ijedudev.2020.102330>
- Aksnes, D. W., Schneider, J. W., & Gunnarsson, M. (2012). Ranking national research systems by citation indicators. A comparative analysis using whole and fractionalised counting methods. *Journal of Informetrics*, 6(1), 36-43. <https://doi.org/10.1016/j.joi.2011.08.002>
- Albarillo, F. (2014). Language in social science databases: English versus non-English articles in JSTOR and Scopus. *Behavioral & Social Sciences Librarian*, 33(2), 77-90. <https://doi.org/10.1080/01639269.2014.904693>
- Albrecht, K., Frost, M., & Handtke, U. (1994, November 2-4). Informetrische Vermessung eines Forschungsinstituts. In W. Rauch, F. Strohmeier, H. Hiller, & C. Schlögl (Eds.), *Mehrwert von information – Professionalisierung der Informationsarbeit: Proceedings des 4. Internationalen Symposiums für Informationswissenschaft (ISI '94)* (pp. 151-163). Universitätsverlag Konstanz. German.
- Aljuaid, H., Iftikhar, R., Ahmad, S., Asif, M., & Tanvir Afzal, M. (2021). Important citation identification using sentiment analysis of in-text citations. *Telematics and Informatics*, 56, 101492. <https://doi.org/10.1016/j.tele.2020.101492>
- Almind, T. C., & Ingwersen, P. (1997). Informetric analyses on the world wide web: Methodological approaches to 'webometrics'. *Journal of Documentation*, 53(4), 404-426. <https://doi.org/10.1108/EUM0000000007205>
- Altvater-Mackensen, N., Balicki, G., Bestakowa, L., Bocatius, B., Braun, J., Brehmer, L., Brune, V., Eigemeier, K., Erdem, F., Fritscher, R., Jacobs, A., Klingsporn, B., Kosinski, M., Kuntze, J., Lee, J. R., Osterhage, A., Probst, M., Risch, T., Schmitt, T., . . . Werner, K. (2005). Science and technology in the region: The output of regional science and technology, its strengths and its leading institutions. *Scientometrics*, 63(3), 463-529. <https://doi.org/10.1007/s11192-005-0225-z>
- Añazco, D., Nicolalde, B., Espinosa, I., Camacho, J., Mushtaq, M., Gimenez, J., & Teran, E. (2021). Publication rate and citation counts for preprints released during the COVID-19 pandemic: the good, the bad and the ugly. *PeerJ*, 9, e10927. <https://doi.org/10.7717/peerj.10927>
- Archambault, É. (2002). Methods of using patents in cross-country comparisons. *Scientometrics*, 54(1), 15-30. <https://doi.org/10.1023/A:1015654903218>
- Auger, P. (1998). *Information sources in grey literature: Guides to information sources* (4th ed.). De Gruyter Saur.
- Aviv-Reuven, S., & Rosenfeld, A. (2021). Publication patterns' changes due to the COVID-19 pandemic: A longitudinal and short-term scientometric analysis. *Scientometrics*, 126(8), 6761-6784. <https://doi.org/10.1007/s11192-021-04059-x>
- Baas, J., Schotten, M., Plume, A., Côté, G., & Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative Science Studies*, 1(1), 377-386. https://doi.org/10.1162/qss_a_00019
- Bar-Ilan, J. (2006). An ego-centric citation analysis of the works of Michael O. Rabin based on multiple citation indexes. *Information Processing & Management*, 42(6), 1553-1566. <https://doi.org/10.1016/j.ipm.2006.03.019>
- Bar-Ilan, J. (2019). Data collection from the web for informetric purposes. In W. Glänzel, H. F. Moed, U. Schmoch, & M. Thelwall (Eds.), *Springer handbook of science and technology indicators* (pp. 781-800). Springer.
- Bar-Ilan, J., & Halevi, G. (2017). Post retraction citations in context: A case study. *Scientometrics*, 113(1), 547-565. <https://doi.org/10.1007/s11192-017-2242-0>
- Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review*, 13(5), 407-424. <https://doi.org/10.1108/eb024320>
- Birkle, C., Pendlebury, D. A., Schnell, J., & Adams, J. (2020). Web of Science as a data source for research on scientific and scholarly activity. *Quantitative Science Studies*, 1(1), 363-376. https://doi.org/10.1162/qss_a_00018
- Borgman, C. L., & Furner, J. (2002). Scholarly communication and bibliometrics. *Annual Review of Information Science and Technology*, 36(1), 2-72. <https://doi.org/10.1002/aris.1440360102>
- Bornmann, L., Wagner, C., & Leydesdorff, L. (2015). BRICS countries and scientific excellence: A bibliometric analysis of most frequently cited papers. *Journal of the Association for Information Science and Technology*, 66(7), 1507-1513. <https://doi.org/10.1002/asi.23333>
- Boxenbaum, H., Pivinski, F., & Ruberg, S. J. (1987). Publication rates of pharmaceutical scientists: Application of the waring distribution. *Drug Metabolism Reviews*, 18(4), 553-571. <https://doi.org/10.3109/03602538708994132>
- Briet, S. (2006). *What is documentation?* (R. E. Day, L. Martinet, & H. G. B. Anghelescu, Trans.). Scarecrow Press.

(Original work published 1951)

- Brin, S., & Page, L. (1998). The anatomy of a large-scale hyper-textual web search engine. *Computer Networks and ISDN Systems*, 30(1-7), 107-117. [https://doi.org/10.1016/S0169-7552\(98\)00110-X](https://doi.org/10.1016/S0169-7552(98)00110-X)
- Broad, W. J. (1981). The publishing game: Getting more for less. *Science*, 211(4487), 1137-1139. <https://doi.org/10.1126/science.7008199>
- Buckland, M. K. (1997). What is a "document"? *Journal of the American Society for Information Science*, 48(9), 804-809. [https://doi.org/10.1002/\(SICI\)1097-4571\(199709\)48:9<804::AID-ASI5>3.0.CO;2-V](https://doi.org/10.1002/(SICI)1097-4571(199709)48:9<804::AID-ASI5>3.0.CO;2-V)
- Cai, L., Tian, J., Liu, J., Bai, X., Lee, I., Kong, X., & Xia, F. (2019). Scholarly impact assessment: A survey of citation weighting solutions. *Scientometrics*, 118(2), 453-478. <https://doi.org/10.1007/s11192-018-2973-6>
- Cambrosio, A., Cointet, J. P., & Abdo, A. H. (2020). Beyond networks: Aligning qualitative and computational science studies. *Quantitative Science Studies*, 1(3), 1017-1024. https://doi.org/10.1162/qss_a_00055
- Carnap, R. (1928). *Der logische Aufbau der Welt*. Meiner. German.
- Carnap, R. (1931). Die physikalische sprache als Universal-sprache der Wissenschaft. *Erkenntnis*, 2(1), 432-465. German. <https://doi.org/10.1007/BF02028172>
- Carnap, R. (1932). Über Protokollsätze. *Erkenntnis*, 3(1), 215-228. <https://doi.org/10.1007/BF01886421>
- Cavaller, V. (2009). Scientometrics and patent bibliometrics in RUL analysis: A new approach to valuation of intangible assets. *Vine*, 39(1), 80-91. <https://doi.org/10.1108/03055720910962461>
- Chavarro, D., Ràfols, I., & Tang, P. (2018). To what extent is inclusion in the Web of Science an indicator of journal 'quality'? *Research Evaluation*, 27(2), 106-118. <https://doi.org/10.1093/reseval/rvy001>
- Chen, X. (2019). Beall's list and Cabell's blacklist: A comparison of two lists of predatory OA journals. *Serials Review*, 45(4), 219-226. <https://doi.org/10.1080/00987913.2019.1694810>
- Chiarelli, A., Johnson, R., Pinfield, S., & Richens, E. (2019). Pre-prints and scholarly communication: Adoption, practices, drivers and barriers. *F1000Research*, 8, 971. <https://doi.org/10.12688/f1000research.19619.2>
- Cobey, K. D., Lalu, M. M., Skidmore, B., Ahmadzai, N., Grudniewicz, A., & Moher, D. (2018). What is a predatory journal? A scoping review. *F1000Research*, 7, 1001. <https://doi.org/10.12688/f1000research.15256.2>
- da Silveira Bueno, C., da Silveira, J. M. F. J., Buainain, A. M., & Poz, M. E. S. D. (2018). Applying an IPC network to identify the bioenergy technological frontier. *Revista Brasileira De Inovação*, 17(2), 259-286. <https://doi.org/10.20396/rbi.v17i2.8652020>
- Del Corso, G. M., & Romani, F. (2009). Versatile weighting strategies for a citation-based research evaluation model. *Bulletin of the Belgian Mathematical Society - Simon Stevin*, 16(4), 723-743. <https://doi.org/10.36045/bbms/1257776244>
- Ding, Y., Zhang, G., Chambers, T., Song, M., Wang, X., & Zhai, C. (2014). Content-based citation analysis: The next generation of citation analysis. *Journal of the Association for Information Science and Technology*, 65(9), 1820-1833. <https://doi.org/10.1002/asi.23256>
- Doerr, M. (2001). Semantic problems of thesaurus mapping. *Journal of Digital Information*, 1(8), 1-15. [https://www.backbonethesaurus.eu/sites/default/files/Semantic-problems-of-thesaurus-mapping\(2001\).pdf](https://www.backbonethesaurus.eu/sites/default/files/Semantic-problems-of-thesaurus-mapping(2001).pdf)
- Docampo, D., & Bessoule, J. J. (2019). A new approach to the analysis and evaluation of the research output of countries and institutions. *Scientometrics*, 119(2), 1207-1225. <https://doi.org/10.1007/s11192-019-03089-w>
- Dorsch, I. (2017). Relative visibility of authors' publications in different information services. *Scientometrics*, 112(2), 917-925. <https://doi.org/10.1007/s11192-017-2416-9>
- Dorsch, I., Askeridis, J., & Stock, W. (2018). Truebounded, overbounded, or underbounded? Scientists' personal publication lists versus lists generated through bibliographic information services. *Publications*, 6(1), 7. <https://doi.org/10.3390/publications6010007>
- Dorsch, I., & Haustein, S. (2023). Bibliometrie. In R. Kuhlen, D. Lewandowski, W. Semar, & C. Womser-Hacker (Eds.), *Grundlagen der Informationswissenschaft* (pp. 271-280). De Gruyter Saur.
- Dorsch, I., Jeffrey, A., Ebrahimzadeh, S., Maggio, L. A., & Haustein, S. (2021, July 12-15). Metrics literacies: On the state of the art of multimedia scholarly metrics education. In W. Glänzel, S. Heeffer, P. S. Chi, & R. Rousseau (Eds.), *Proceedings of the 18th International Conference on Scientometrics & Informetrics (ISSI2021)* (pp. 1465-1466). ISSI.
- Egghe, L., Guns, R., & Rousseau, R. (2013). Measuring co-authors' contribution to an article's visibility. *Scientometrics*, 95(1), 55-67. <https://doi.org/10.1007/s11192-012-0832-4>
- Egghe, L., & Rousseau, R. (1990). *Introduction to informetrics: Quantitative methods in library, documentation and information science*. Elsevier.
- El-Ouahi, J. (2022). The Arabic Citation Index - Toward a better understanding of Arab scientific literature. *Research Square*. <https://doi.org/10.21203/rs.3.rs-1917781/v2>
- Erdt, M., Nagarajan, A., Sin, S. C. J., & Theng, Y. L. (2016). Altmetrics: An analysis of the state-of-the-art in measuring

- research impact on social media. *Scientometrics*, 109(2), 1117-1166. <https://doi.org/10.1007/s11192-016-2077-0>
- Farid, F. O. (2021). On a new index for research assessment. *Journal of Information Science Theory and Practice*, 9(3), 56-75. <https://doi.org/10.1633/JISTaP.2021.9.3.5>
- Fasce, A. (2017). What do we mean when we speak of pseudo-science? The development of a demarcation criterion based on the analysis of twenty-one previous attempts. *Disputatio. Philosophical Research Bulletin*, 6(7), 459-488. <https://doi.org/10.5281/zenodo.1433737>
- Faust, K. (1992, March 19-20). Patentanmeldungen als Indikator von technologischen Entwicklungen – Ergebnisse und Erfahrungen mit der IFO Patentstatistik. In A. Editor & B. Editor (Eds.), *Proceedings of the 1st EPO-IFO-Workshop: Results and Methods of Economic Patent Research* (pp. 155-172). European Patent Office. German.
- Fernandez-Beanato, D. (2020). The multicriterial approach to the problem of demarcation. *Journal for General Philosophy of Science*, 51(3), 375-390. <https://doi.org/10.1007/s10838-020-09507-5>
- Feyerabend, P. (1975). *Against method: Outline of an anarchistic theory of knowledge*. New Left Books.
- Finkenstaedt, T. (1986). Forschungsmessung in den Geisteswissenschaften. In R. Fisch, & H. D. Daniel (Eds.), *Messung und Förderung von Forschungsleistung* (pp. 51-60). Universitätsverlag Konstanz. German.
- Flanagin, A., Carey, L. A., Fontanarosa, P. B., Phillips, S. G., Pace, B. P., Lundberg, G. D., & Rennie, D. (1998). Prevalence of articles with honorary authors and ghost authors in peer-reviewed medical journals. *Journal of the American Medical Association*, 280(3), 222-224. <https://doi.org/10.1001/jama.280.3.222>
- Fraser, N., Brierley, L., Dey, G., Polka, J. K., Pálffy, M., Nanni, F., & Coates, J. A. (2021). The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape. *PLoS Biology*, 19(4), e3000959. <https://doi.org/10.1371/journal.pbio.3000959>
- Fraser, N., Mayr, P., & Peters, I. (2022). Motivations, concerns and selection biases when posting preprints: A survey of bioRxiv authors. *PLoS One*, 17(11), e0274441. <https://doi.org/10.1371/journal.pone.0274441>
- Fraser, N., Momeni, F., Mayr, P., & Peters, I. (2020). The relationship between bioRxiv preprints, citations and altmetrics. *Quantitative Science Studies*, 1(2), 618-638. https://doi.org/10.1162/qss_a_00043
- Fröhlich, G. (1999). Das messen des leicht meßbaren: Output-indikatoren, impact-maße: Artefakte der szientometrie? In J. Becker, & W. Göhring (Eds.), *Kommunikation statt markt: Zu einer alternativen theorie der informationsgesellschaft* (pp. 27-38). GMD - Forschungszentrum Informationstechnik.
- García-Villar, C. (2021). A critical review on altmetrics: Can we measure the social impact factor? *Insights Imaging*, 12, 92. <https://doi.org/10.1186/s13244-021-01033-2>
- Garfield, E. (1979). *Citation indexing - Its theory and application in science, technology, and humanities*. Wiley.
- Garfield, E., Paris, S. W., & Stock, W. G. (2006). HistCite™: A software tool for informetric analysis of citation linkage. *NFD Information-Wissenschaft und Praxis*, 57(8), 391-400. https://www.isi.hhu.de/fileadmin/redaktion/Fakultaeten/Philosophische_Fakultaet/Sprache_und_Information/Informationswissenschaft/Dateien/Wolfgang_G._Stock/Daten_vor_2011/1166781999eugenegarlf.pdf
- Gaufriau, M. (2017). A categorization of arguments for counting methods for publication and citation indicators. *Journal of Informetrics*, 11(3), 672-684. <https://doi.org/10.1016/j.joi.2017.05.009>
- Gaufriau, M. (2021). Counting methods introduced into the bibliometric research literature 1970–2018: A review. *Quantitative Science Studies*, 2(3), 932-975. https://doi.org/10.1162/qss_a_00141
- Gaufriau, M., & Larsen, P. O. (2005). Counting methods are decisive for rankings based on publication and citation studies. *Scientometrics*, 64(1), 85-93. <https://doi.org/10.1007/s11192-005-0239-6>
- Gaufriau, M., Larsen, P. O., Maye, I., Roulin-Perriard, A., & von Ins, M. (2007). Publication, cooperation and productivity measures in scientific research. *Scientometrics*, 73(2), 175-214. <https://doi.org/10.1007/s11192-007-1800-2>
- Gaufriau, M., Larsen, P. O., Maye, I., Roulin-Perriard, A., & von Ins, M. (2008). Comparisons of results of publication counting using different methods. *Scientometrics*, 77(1), 147-176 (2008). <https://doi.org/10.1007/s11192-007-1934-2>
- Glänzel, W. (2008). Seven myths in bibliometrics about facts and fiction in quantitative science studies. *COLLNET Journal of Scientometrics and Information Management*, 2(1), 9-17. <https://doi.org/10.1080/09737766.2008.10700836>
- Grünbaum, A. (1977). Is psychoanalysis a pseudo-science? Karl Popper versus Sigmund Freud. *Zeitschrift Für Philosophische Forschung*, 31(3), 333-353. <http://www.jstor.org/stable/20482837>
- Gremm, J., Barth, J., Fietkiewicz, K. J., & Stock, W. G. (2018). *Transitioning towards a knowledge society: Qatar as a case study*. Springer.
- Halevi, G., Moed, H., & Bar-Ilan, J. (2017). Suitability of Google Scholar as a source of scientific information and as a source of data for scientific evaluation—Review of the literature. *Journal of Informetrics*, 11(3), 823-834. <https://doi.org/10.1007/s11192-016-2077-0>

- org/10.1016/j.joi.2017.06.005
- Hansson, S. O. (2013). Defining pseudoscience and science. In M. Pigliucci, & M. Boudry (Eds.), *Philosophy of pseudoscience: Reconsidering the demarcation problem* (pp. 61-78). University of Chicago Press.
- Haustein, S. (2012). *Multidimensional journal evaluation: Analyzing scientific periodicals beyond the impact factor*. De Gruyter Saur.
- Haustein, S. (2016). Grand challenges in altmetrics: Heterogeneity, data quality and dependencies. *Scientometrics*, 108(1), 413-423. <https://doi.org/10.1007/s11192-016-1910-9>
- Haustein, S. (2019). Scholarly Twitter metrics. In W. Glänzel, H. F. Moed, U. Schmoch, & M. Thelwall (Eds.), *Springer handbook of science and technology indicators* (pp. 729-760). Springer.
- Henderson, M. (2010). Problems with peer review. *British Medical Journal*, 340, c1409. <https://doi.org/10.1136/bmj.c1409>
- Henrichs, N. (1994). Begriffswandel in Datenbanken: Kontextuelle Inhaltsanalyse für Disambiguierung und ideengeschichtliche Analyse. In H. Best, B. Endres-Niggemeyer, M. Herfurth, & H. P. Ohly (Eds.), *Informations- und Wissensverarbeitung in den Sozialwissenschaften* (pp. 225-239). VS Verlag für Sozialwissenschaften. German.
- Herb, U. (2016). Altmetrics zwischen Revolution und Dienstleistung: Eine methodische und konzeptionelle Kritik. In H. Staubmann (Ed.), *Soziologie in Österreich – Internationale Verflechtungen* (pp. 387-410). Innsbruck University Press. German.
- Herzog, C., Hook, D., & Konkiel, S. (2020). Dimensions: Bringing down barriers between scientometricians and data. *Quantitative Science Studies*, 1(1), 387-395. https://doi.org/10.1162/qss_a_00020
- Heumüller, R., Nielebock, S., Krüger, J., & Ortmeier, F. (2020). Publish or perish, but do not forget your software artifacts. *Empirical Software Engineering*, 25(6), 4585-4616. <https://doi.org/10.1007/s10664-020-09851-6>
- Hilbert, F., Barth, J., Gremm, J., Gros, D., Haiter, J., Henkel, M., Reinhardt, W., & Stock, W. G. (2015). Coverage of academic citation databases compared with coverage of scientific social media: Personal publication lists as calibration parameters. *Online Information Review*, 39(2), 255-264. <https://doi.org/10.1108/OIR-07-2014-0159>
- Holcombe, A. (2019). Farewell authors, hello contributors. *Nature*, 571(7764), 147. <https://doi.org/10.1038/d41586-019-02084-8>
- Holmberg, K. (2015). *Altmetrics for information professionals: Past, present and future*. Chandos Publishing.
- Honka, A., Orszulok, L., Dorsch, I., & Frommelius, N. (2015, May 19-21). Topical impact analysis - A new informetric indicator for the assessment of a scientific institution. In F. Pehar, C. Schlögl, & C. Wolff (Eds.), *Proceedings of the 14th International Symposium on Information Science (ISI 2015): Re:inventing Information Science in the Networked Society* (pp. 410-421). Verlag Werner Hülsbusch.
- Hu, G., Wang, L., Ni, R., & Liu, W. (2020). Which h-index? An exploration within the Web of Science. *Scientometrics*, 123(3), 1225-1233. <https://doi.org/10.1007/s11192-020-03425-5>
- Huang, C. K., Neylon, C., Kenworthy, C. B., Hosking, R., Montgomery, L., Wilson, K., & Ozaygen, A. (2020). Comparison of bibliographic data sources: Implications for the robustness of university rankings. *Quantitative Science Studies*, 1(2), 445-478. https://doi.org/10.1162/qss_a_00031
- Huang, M. H., Lin, C. S., & Chen, D. Z. (2011). Counting methods, country rank changes, and counting inflation in the assessment of national research productivity and impact. *Journal of the American Society for Information Science and Technology*, 62(12), 2427-2436. <https://doi.org/10.1002/asi.21625>
- IFLA Study Group on the Functional Requirements for Bibliographic Records. (1998). *Functional requirements for bibliographic records*. K. G. Saur.
- Ingwersen, P. (2002, July 21-25). Cognitive perspectives of document representation. In H. Bruce, R. Fidel, P. Ingwersen, & P. Vakkari (Eds.), *Proceedings of the 4th International Conference on Conceptions of Library and Information Science (CoLIS 4)* (pp. 285-300). Libraries Unlimited.
- Järvelin, K., & Vakkari, P. (2022). LIS research across 50 years: Content analysis of journal articles. *Journal of Documentation*, 78(7), 65-88. <https://doi.org/10.1108/JD-03-2021-0062>
- Jansz, C. N. M., & le Pair, C. (1991). Die bibliometrische Unsichtbarkeit des technischen Fortschritts. In P. Weingart, R. Sehringer, & M. Winterhager (Eds.), *Indikatoren der Wissenschaft und Technik: Theorie, Methoden, Anwendungen* (pp. 209-223). Campus-Verlag. German.
- Jesenko, B., & Schlögl, C. (2021). The effect of web of science subject categories on clustering: The case of data-driven methods in business and economic sciences. *Scientometrics*, 126(8), 6785-6801. <https://doi.org/10.1007/s11192-021-04060-4>
- Jin, B., & Wang, B. (1999). Chinese science citation database: Its construction and application. *Scientometrics*, 45(2), 325-332. <https://doi.org/10.1007/BF02458440>
- Johansson, M. A., Reich, N. G., Meyers, L. A., & Lipsitch, M. (2018). Preprints: An underutilized mechanism to accelerate outbreak science. *PLoS Medicine*, 15(4), e1002549.

- <https://doi.org/10.1371/journal.pmed.1002549>
- Kang, D., & Evans, J. (2020). Against method: Exploding the boundary between qualitative and quantitative studies of science. *Quantitative Science Studies*, 1(3), 930-944. https://doi.org/10.1162/qss_a_00056
- Kamrani, P., Dorsch, I., & Stock, W. G. (2020). Publikationen, Zitationen und H-index im Meinungsbild Deutscher Universitätsprofessoren. *Beiträge zur Hochschulforschung*, 42(3), 78-98. German.
- Kamrani, P., Dorsch, I., & Stock, W. G. (2021). Do researchers know what the h-index is? And how do they estimate its importance? *Scientometrics*, 126(7), 5489-5508. <https://doi.org/10.1007/s11192-021-03968-1>
- Katz, E., Blumler, J. G., & Gurevitch, M. (1973-1974). Uses and gratifications research. *Public Opinion Quarterly*, 37(4), 509-523. <https://doi.org/10.1086/268109>
- Kassian, A., & Melikhova, L. (2019). Russian Science Citation Index on the WoS platform: A critical assessment. *Journal of Documentation*, 75(5), 1162-1168. <https://doi.org/10.1108/JD-02-2019-0033>
- Keng, S. T., Ooi, P. B., Pang, K. Y., & Phan, C. W. (2022). Awareness, familiarity and usage of traditional research metrics and altmetrics among academics in Malaysia. *Malaysian Journal of Library and Information Science*, 27(3), 1-19. <https://doi.org/10.22452/mjlis.vol27no3.1>
- Kleminski, R., Kazienko, P., & Kajdanowicz, T. (2022). Analysis of direct citation, co-citation and bibliographic coupling in scientific topic identification. *Journal of Information Science*, 48(3), 349-373. <https://doi.org/10.1177/0165551520962775>
- Ko, Y. M., Cho, S. R., & Park, Y. S. (2011). A study on the optimization of KCI-based index (Kor-Factor) in evaluating Korean journals. *Scientometrics*, 88(1), 61-71. <https://doi.org/10.1007/s11192-011-0384-z>
- Kostoff, R. N., & Martinez, W. L. (2005). Is citation normalization realistic? *Journal of Information Science*, 31(1), 57-61. <https://doi.org/10.1177/0165551505049260>
- Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). University of Chicago Press.
- Kutlača, D., Babić, D., Živković, L., & Štrbac, D. (2015). Analysis of quantitative and qualitative indicators of SEE countries scientific output. *Scientometrics*, 102(1), 247-265. <https://doi.org/10.1007/s11192-014-1290-y>
- Lakatos, I. (1974). Science and pseudoscience. *Conceptus*, 8, 5-9.
- Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Bibliometrics: Global gender disparities in science. *Nature*, 504(7479), 211-213. <https://doi.org/10.1038/504211a>
- Larivière, V., Pontille, D., & Sugimoto, C. R. (2021). Investigating the division of scientific labor using the Contributor Roles Taxonomy (CRediT). *Quantitative Science Studies*, 2(1), 111-128. doi: https://doi.org/10.1162/qss_a_00097
- Lasswell, H. D. (1948). The structure and function of communication in society. In L. Bryson (Ed.), *The communication of ideas* (pp. 37-51). Harper & Brothers.
- Lee, C. J., Sugimoto, C. R., Zhang, G., & Cronin, B. (2013). Bias in peer review. *Journal of the American Society for Information Science and Technology*, 64(1), 2-17. <https://doi.org/10.1002/asi.22784>
- Leydesdorff, L. (1995). *The challenge of scientometrics: The development, measurement, and self-organization of scientific communications*. DSWO Press.
- Leydesdorff, L., Ràfols, I., & Milojević, S. (2020). Bridging the divide between qualitative and quantitative science studies. *Quantitative Science Studies*, 1(3), 918-926. https://doi.org/10.1162/qss_e_00061
- Leydesdorff, L., & Shin, J. C. (2011). How to evaluate universities in terms of their relative citation impacts: Fractional counting of citations and the normalization of differences among disciplines. *Journal of the American Society for Information Science and Technology*, 62(6), 1146-1155. <https://doi.org/10.1002/asi.21511>
- Leydesdorff, L., & Welbers, K. (2011). The semantic mapping of words and co-words in contexts. *Journal of Informetrics*, 5(3), 469-475. <https://doi.org/10.1016/j.joi.2011.01.008>
- Li, M. (2018). Classifying and ranking topic terms based on a novel approach: Role differentiation of author keywords. *Scientometrics*, 116(1), 77-100. <https://doi.org/10.1007/s11192-018-2741-7>
- Lin, J., & Fenner, M. (2013). Altmetrics in evolution: Defining & redefining the ontology of article-level metrics. *Information Standards Quarterly*, 25(2), 20-26. <http://doi.org/10.3789/isqv25no2.2013.04>
- Liu, J., & Adie, E. (2013). Five challenges in altmetrics: A tool-maker's perspective. *Bulletin of the American Society for Information Science and Technology*, 39(4), 31-34. <https://doi.org/10.1002/bult.2013.1720390410>
- Liu, W. (2019). The data source of this study is Web of Science Core Collection? Not enough. *Scientometrics*, 121(3), 1815-1824. <https://doi.org/10.1007/s11192-019-03238-1>
- Lundberg, G. D. (1993). Statement by the International Committee of Medical Journal Editors on duplicate or redundant publication. *Journal of the American Medical Association*, 270(20), 2495. <https://doi.org/10.1001/jama.1993.03510200101040>
- Ma, N., Guan, J., & Zhao, Y. (2008). Bringing PageRank to the citation analysis. *Information Processing & Management*,

- 44(2), 800-810. <https://doi.org/10.1016/j.ipm.2007.06.006>
- MacRoberts, M. H., & MacRoberts, B. R. (1989). Problems of citation analysis: A critical review. *Journal of the American Society for Information Science*, 40(5), 342-349. [https://doi.org/10.1002/\(SICI\)1097-4571\(198909\)40:5%3C342::AID-ASI7%3E3.0.CO;2-U](https://doi.org/10.1002/(SICI)1097-4571(198909)40:5%3C342::AID-ASI7%3E3.0.CO;2-U)
- MacRoberts, M. H., & MacRoberts, B. R. (2018). The mismeasure of science: Citation analysis. *Journal of the Association for Information Science and Technology*, 69(3), 474-482. <https://doi.org/10.1002/asi.23970>
- Maron, M. E. (1977). On indexing, retrieval and the meaning of about. *Journal of the American Society for Information Science*, 28(1), 38-43. <https://doi.org/10.1002/asi.4630280107>
- Marsh, R. M. (2015). The role of institutional repositories in developing the communication of scholarly research. *OCLC Systems & Services: International Digital Library Perspectives*, 31(4), 163-195. <https://doi.org/10.1108/OCLC-04-2014-0022>
- Martín-Martín, A., Thelwall, M., Orduna-Malea, E., & López-Cózar, E. D. (2021). Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: A multidisciplinary comparison of coverage via citations. *Scientometrics*, 126(1), 871-906. <https://doi.org/10.1007/s11192-020-03690-4>
- McKiernan, G. (2000). arXiv.org: The Los Alamos National Laboratory e-print server. *International Journal on Grey Literature*, 1(3), 127-138. <https://doi.org/10.1108/14666180010345564>
- Martyn, C. (2005). The malign influence of impact factors. *Journal of the Royal Society of Medicine*, 98(12), 556. <https://doi.org/10.1177/014107680509801221>
- Massarani, L., & Neves, L. F. F. (2022). Reporting COVID-19 preprints: Fast science in newspapers in the United States, the United Kingdom and Brazil. *Ciencia & Saude Coletiva*, 27(3), 957-968. <https://doi.org/10.1590/1413-81232022273.20512021>
- Meschede, C., & Siebenlist, T. (2018). Cross-metric compatibility and inconsistencies of altmetrics. *Scientometrics*, 115(1), 283-297. <https://doi.org/10.1007/s11192-018-2674-1>
- Migheli, M., & Ramello, G. B. (2021). The unbearable lightness of scientometric indices. *Managerial and Decision Economics*, 42(8), 1933-1944. <https://doi.org/10.1002/mde.3486>
- Mikhailov, O. V. (2021). Introduction from the guest editor of special issue "Modern problems of scientometric assessment of publication activity". *Publications*, 9(2), 19. <https://doi.org/10.3390/publications9020019>
- Mingers, J., & Leydesdorff, L. (2015). A review of theory and practice in scientometrics. *European Journal of Operational Research*, 246(1), 1-19. <https://doi.org/10.1016/j.ejor.2015.04.002>
- Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 106(1), 213-228. <https://doi.org/10.1007/s11192-015-1765-5>
- Moussa, S. (2021). Citation contagion: A citation analysis of selected predatory marketing journals. *Scientometrics*, 126(1), 485-506. <https://doi.org/10.1007/s11192-020-03729-6>
- Mousseau, M. C. (2003). Parapsychology: Science or pseudo-science? *Journal of Scientific Exploration*, 17(2), 271-282. https://www.scientificexploration.org/docs/17/jse_17_2_mousseau.pdf
- Narin, F., Olivastro, D., & Stevens, K. A. (1994). Bibliometrics/ theory, practice and problems. *Evaluation Review*, 18(1), 65-76. <https://doi.org/10.1177/0193841X9401800107>
- Neuhäusler, P., & Frietsch, R. (2013). Patent families as macro level patent value indicators: Applying weights to account for market differences. *Scientometrics*, 96(1), 27-49. <https://doi.org/10.1007/s11192-012-0870-y>
- Neurath, O. (1932). Protokollsätze. *Erkenntnis*, 3(1), 204-214. German. <https://doi.org/10.1007/BF01886420>
- Ni, C., Smith, E., Yuan, H., Larivière, V., & Sugimoto, C. R. (2021). The gendered nature of authorship. *Science Advances*, 7(36), eabe4639. <https://www.science.org/doi/10.1126/sciadv.abe4639>
- Ninkov, A., Gregory, K., Peters, I., & Haustein, S. (2021, April 30). Datasets on DataCite - An initial bibliometric investigation. *Proceeding of the 18th International Conference on Scientometrics & Informetrics (ISSI 2021)* (pp. 1-6). Zenodo.
- Organisation for Economic Co-operation and Development (OECD). (2015). *Frascati manual 2015: Guidelines for collecting and reporting data on research and experimental development*. OECD.
- Olcay, G. A., & Bulu, M. (2017). Is measuring the knowledge creation of universities possible?: A review of university rankings. *Technological Forecasting and Social Change*, 123, 153-160. <https://doi.org/10.1016/j.techfore.2016.03.029>
- Osório, A. (2018). On the impossibility of a perfect counting method to allocate the credits of multi-authored publications. *Scientometrics*, 116(3), 2161-2173. <https://doi.org/10.1007/s11192-018-2815-6>
- Packer, A. L. (2020). The pasts, presents, and futures of SciELO. In M. P. Eve, & J. Gray (Eds.), *Reassembling scholarly communications: Histories, infrastructures, and global politics of open access* (pp. 297-313). MIT Press.
- Peters, I., Kraker, P., Lex, E., Gumpenberger, C., & Gorraiz, J. (2016). Research data explored: An extended analysis of citations and altmetrics. *Scientometrics*, 107(2), 723-744.

- <https://doi.org/10.1007/s11192-016-1887-4>
- Peters, I., Kraker, P., Lex, E., Gumpenberger, C., & Gorraiz, J. I. (2017). Zenodo in the spotlight of traditional and new metrics. *Frontiers in Research Metrics and Analytics*, 2, 13. <https://doi.org/10.3389/frma.2017.00013>
- Poppema, S. (2020). Citations are the currency of science: Then there are the counterfeiters. *The Scientist*, 34(12), 14-15. <https://www.the-scientist.com/critic-at-large/citations-are-the-currency-of-science-68204>
- Popper, K. R. (1935). *Logik der Forschung: Zur Erkenntnistheorie der Modernen Naturwissenschaft*. Julius Springer Verlag. German.
- Popper, K. R. (1962). *Conjectures and refutations: The growth of scientific knowledge*. Basic Books.
- Prancutè, R. (2021). Web of Science (WoS) and Scopus: The titans of bibliographic information in today's academic world. *Publications*, 9(1), 12. <https://doi.org/10.3390/publications9010012>
- Priem, J. (2010). *I like the term #articlelevelmetrics, but it fails to imply *diversity* of measures. Lately, I'm liking #altmetrics*. <https://twitter.com/jasonpriem/status/25844968813>
- Priem, J. (2014). Altmetrics. In B. Cronin, & C. R. Sugimoto (Eds.), *Beyond bibliometrics: Harnessing multidimensional indicators of scholarly impact* (pp. 263-287). MIT Press.
- Priem, J., Piwowar, H., & Orr, R. (2022). OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. *arXiv*. <https://doi.org/10.48550/arXiv.2205.01833>
- Quay, P. M. (1974). Progress as a demarcation criterion for the sciences. *Philosophy of Science*, 41(2), 154-170. <http://www.jstor.org/stable/186865>
- Rahman, M. T., Regenstein, J. M., Kassim, N. L. A., & Haque, N. (2017). The need to quantify authors' relative intellectual contributions in a multi-author paper. *Journal of Informetrics*, 11(1), 275-281. <https://doi.org/10.1016/j.joi.2017.01.002>
- Raman, V. V. (2001). Science and religion: Some demarcation criteria. *Zygon*, 36(3), 541-556. <https://doi.org/10.1111/0591-2385.00380>
- Rauter, J. (2005, September 13-16). Wissenschaftliche Forschung und ihre Evaluation: Expansive Zitationsanalyse auf deskriptiv-intertextueller Basis. In E. Pipp (Ed.), *Proceedings of the 11th Austrian Online Information Meeting, 12th Austrian Documentary Day (ODOK '05)* (pp. 160-184). Wolfgang Neugebauer Verlag. German.
- Refinetti, R. (1990). In defense of the least publishable unit. *FASEB Journal*, 4(1), 128-129. <https://doi.org/10.1096/fasebj.4.1.2295373>
- Reichmann, G. (2020). Was zählt eine Publikation? Möglichkeiten zur Durchführung einer Forschungsevaluierung im Bereich der Betriebswirtschaftslehre. In I. Dorsch, K. Fietkiewicz, A. Ilhan, C. Meschede, & T. Siebenlist (Eds.), *Facetten von Wolf Stock und ihre Bedeutung für die Informationswissenschaft: Festschrift zu Ehren von Wolfgang G. Stock, Schriften zur Informationswissenschaft* (pp. 15-34). Verlag Werner Hülsbusch. German.
- Reichmann, G., & Schlögl, C. (2021). Möglichkeiten zur Steuerung der Ergebnisse einer Forschungsevaluierung: Ein Vergleich der Forschungsleistung zweier informationswissenschaftlicher Institute. *Information - Wissenschaft & Praxis*, 72(4), 212-220. German. <https://doi.org/10.1515/iwp-2021-2148>
- Reichmann, G., & Schlögl, C. (2022). On the possibilities of presenting the research performance of an institute over a long period of time: The case of the Institute of Information Science at the University of Graz in Austria. *Scientometrics*, 127(6), 3193-3223. <https://doi.org/10.1007/s11192-022-04377-8>
- Reichmann, G., Schlögl, C., Stock, W. G., & Dorsch, I. (2022). Forschungsevaluierung auf Institutsebene – Der Einfluss der gewählten Methodik auf die Ergebnisse. *Beiträge zur Hochschulforschung*, 44(1), 74-97. German. https://www.bzh.bayern.de/fileadmin/user_upload/Publikationen/Beitraege_zur_Hochschulforschung/2022/2022-1-Reichmann-Schloegl-Stock-Dorsch.pdf
- Riegelman, A. (2022). Research square. *Charleston Advisor*, 24(2), 43-45. <https://doi.org/10.5260/chara.24.2.43>
- Robinson-García, N., Jiménez-Contreras, E., & Torres-Salinas, D. (2016). Analyzing data citation practices using the data citation index. *Journal of the Association for Information Science and Technology*, 67(12), 2964-2975. <https://doi.org/10.1002/asi.23529>
- Roth, J. (1981). Least publishable unit. *Science*, 212(4494), 494. <https://doi.org/10.1126/science.212.4494.494.b>
- Röbken, H. (2011). Forschungsproduktivität von Wissenschaftlern und Wissenschaftlerinnen – Eine empirische Analyse von Publikationsaktivitäten vor und nach der Berufung. *Beiträge zur Hochschulforschung*, 33(3), 62-81. German. https://www.bzh.bayern.de/fileadmin/news_import/3-2011-roebken_01.pdf
- Rousseau, R., & Egghe, L., & Guns, R. (2018). *Becoming metric-wise: A bibliometric guide for researchers*. Chandos.
- Scanlan, B. D. (1988). Coverage by current contents and the validity of impact factors. *Serials Librarian*, 13(2-3), 57-66. https://doi.org/10.1300/J123v13n02_07
- Schlögl, C. (2013). Internationale Sichtbarkeit der europäischen und insbesondere der deutschsprachigen Informationswissenschaft [International visibility of the European information science].

- tion science, especially in German language]. *Information - Wissenschaft & Praxis*, 64(1), 1-8. German. <https://doi.org/10.1515/iwp-2013-0001>
- Schlögl, C., Gorraiz, J., Gumpenberger, C., Jack, K., & Kraker, P. (2014). Comparison of downloads, citations and readership data for two information systems journals. *Scientometrics*, 101(2), 1113-1128. <https://doi.org/10.1007/s11192-014-1365-9>
- Schlögl, C., & Stock, W. G. (2008). Practitioners and academics as authors and readers: The case of LIS journals. *Journal of Documentation*, 64(5), 643-666. <https://doi.org/10.1108/00220410810899691>
- Schneijderberg, C. (2018). *Promovieren in den Sozialwissenschaften: Eine sozialisationstheoretische Erschließung des Forschungsfeldes Promotion*. Springer. German.
- Schrader, U., & Hennig-Thurau, T. (2009). VHB-JOURQUAL2: Method, results, and implications of the German Academic Association for Business Research's journal ranking. *BuR Business Research Journal*, 2(2), 180-204. <https://ssrn.com/abstract=1568651>
- Schubert, A., Glänzel, W., & Thijs, B. (2006). The weight of author self-citations. A fractional approach to self-citation counting. *Scientometrics*, 67(3), 503-514. <https://doi.org/10.1556/Scient.67.2006.3.11>
- Seglen, P. O. (1992). The skewness of science. *Journal of the American Society for Information Science*, 43(9), 628-638. [https://doi.org/10.1002/\(SICI\)1097-4571\(199210\)43:9%3C628::AID-ASI5%3E3.0.CO;2-0](https://doi.org/10.1002/(SICI)1097-4571(199210)43:9%3C628::AID-ASI5%3E3.0.CO;2-0)
- Seglen, P. O. (1997). Why the impact factor of journals should not be used for evaluating research. *BMJ*, 314(7079), 498-502. <https://doi.org/10.1136/bmj.314.7079.497>
- Shema, H., Hahn, O., Mazarakis, A., & Peters, I. (2019). Retractions from altmetric and bibliometric perspectives. *Information - Wissenschaft & Praxis*, 70(2-3), 98-110. <https://doi.org/10.1515/iwp-2019-2006>
- Shen, H. W., & Barabási, A. L. (2014). Collective credit allocation in science. *Proceedings of the National Academy of Sciences*, 111(34), 12325-12330. <https://doi.org/10.1073/pnas.1401921111>
- Shu, F., Qiu, J., & Larivière, V. (2021). Mapping the biomedical sciences using Medical Subject Headings: A comparison between MeSH co-assignments and MeSH citation pairs. *Journal of the Medical Library Association*, 109(3), 441-449. <https://doi.org/10.5195/jmla.2021.1173>
- Sile, L., Pölönen, J., Sivertsen, G., Guns, R., Engels, T. C. E., Arefiev, P., Dušková, M., Faurbæk, L., Holl, A., Kulczycki, E., Macan, B., Nelhans, G., Petr, M., Pisk, M., Soós, S., Stojanovski, J., Stone, A., Šušol, J., & Teitelbaum, R. (2018). Comprehensiveness of national bibliographic databases for social sciences and humanities: Findings from a European survey. *Research Evaluation*, 27(4), 310-322. <https://doi.org/10.1093/reseval/rvy016>
- Sjögårde, P., & Didegah, F. (2022). The association between topic growth and citation impact of research publications. *Scientometrics*, 127(4), 1903-1921. <https://doi.org/10.1007/s11192-022-04293-x>
- Smith, R. (2006). Peer review: A flawed process at the heart of science and journals. *Journal of the Royal Society of Medicine*, 99(4), 178-182. <https://doi.org/10.1177/014107680609900414>
- Stock, M., & Stock, W. G. (2006). Intellectual property information: A comparative analysis of main information providers. *Journal of the American Society for Information Science and Technology*, 57(13), 1794-1803. <https://doi.org/10.1002/asi.20498>
- Stock, W. G. (1984). Informetrische Untersuchungsmethoden auf der Grundlage der Textwortmethode. *International Classification*, 11(1984), 151-157. German. <https://www.phil-fak.uni-duesseldorf.de/publikationen/publikationen/publikationen-1990-und-frueher/019-informetrische-untersuchungsmethoden-auf-der-grundlage-der-textwort-methode/>
- Stock, W. G. (1985). Die Bedeutung der Zitatelanalyse für die Wissenschaftsforschung. *Zeitschrift für Allgemeine Wissenschaftstheorie*, 16(2), 304-314. German. <https://doi.org/10.1007/BF01803677>
- Stock, W. G. (1990). Themenanalytische informetrische Methoden: Vorbemerkungen zur Methode. In M. Stock, & W. G. Stock (Eds.), *Psychologie und Philosophie der Grazer Schule* (pp. 7-31). Rodopi. German.
- Stock, W. G. (1994). *Wissenschaftsevaluation: Die Bewertung wissenschaftlicher Forschung und Lehre*. IFO Institut für Wirtschaftsforschung. German.
- Stock, W. G. (2000). Was ist eine Publikation? Zum Problem der Einheitenbildung in der Wissenschaftsforschung. In K. Fuchs-Kittowski, H. Laitko, H. Parthey, & W. Umstätter (Eds.), *Wissenschaft und Digitale Bibliothek: Wissenschaftsforschung Jahrbuch 1998* (pp. 239-282). Gesellschaft für Wissenschaftsforschung. German.
- Stock, W. G. (2001). *Publikation und Zitat: Die problematische Basis empirischer Wissenschaftsforschung* (Band 29). Fachhochschule Köln. German.
- Stock, W. G. (2021). Die Teildatenbanken von Web of Science und deren Subskriptionen von Wissenschaftseinrichtungen: Auswirkungen auf Forschungsevaluation und Sichtbarkeit von Forschenden. *Zeitschrift für Bibliothekswesen und Bibliographie*, 68(3), 152-157. German. <https://doi.org/10.3196/186429502068331>

- Stock, W. G., Dorsch, I., Reichmann, G., & Schlögl, C. (2023). Labor productivity, labor impact, and co-authorship of research institutions: Publications and citations per full-time equivalents. *Scientometrics*, 128(1), 363-377. <https://doi.org/10.1007/s11192-022-04582-5>
- Stock, W. G., & Weber, S. (2006). Facets of informetrics. *Information - Wissenschaft & Praxis*, 57(8), 385-389. https://www.ling.hhu.de/fileadmin/redaktion/Fakultaeten/Philosophische_Fakultaet/Sprache_und_Information/Informationswissenschaft/Dateien/Wolfgang_G._Stock/Daten_vor_2011/1166781846sonjaweber.pdf
- Stock, W. G., & Stock, M. (2013). *Handbook of information science*. De Gruyter Saur.
- Strcic, J., Civljak, A., Glozinic, T., Pacheco, R. L., Brkovic, T., & Puljak, L. (2022). Open data and data sharing in articles about COVID-19 published in preprint servers medRxiv and bioRxiv. *Scientometrics*, 127(5), 2791-2802. <https://doi.org/10.1007/s11192-022-04346-1>
- Strinzel, M., Severin, A., Milzow, K., & Egger, M. (2019). Blacklists and whitelists to tackle predatory publishing: A cross-sectional comparison and thematic analysis. *mBio*, 10(3), e00411-e00419. <https://doi.org/10.1128/mBio.00411-19>
- Sugimoto, C. R., Work, S., Larivière, V., & Haustein, S. (2017). Scholarly use of social media and altmetrics: A review of the literature. *Journal of the Association for Information Science and Technology*, 68(9), 2037-2062. <https://doi.org/10.1002/asi.23833>
- Szomszor, M., & Adie, E. (2022). Overton: A bibliometric database of policy document citations. *Quantitative Science Studies*, 3(3), 624-650. https://doi.org/10.1162/qss_a_00204
- Tague-Sutcliffe, J. (1992). An introduction to informetrics. *Information Processing & Management*, 28(1), 1-3. [https://doi.org/10.1016/0306-4573\(92\)90087-G](https://doi.org/10.1016/0306-4573(92)90087-G)
- Teixeira da Silva, J. A., & Dobránszki, J. (2016). Multiple authorship in scientific manuscripts: Ethical challenges, ghost and guest/gift authorship, and the cultural/disciplinary perspective. *Science and Engineering Ethics*, 22(5), 1457-1472. <https://doi.org/10.1007/s11948-015-9716-3>
- Tennant, J., Bauin, S., James, S., & Kant, J. (2018). The evolving preprint landscape: Introductory report for the knowledge exchange working group on preprints. *MetaArXiv*. <https://doi.org/10.31222/osf.io/796tu>
- Thagard, P. R. (1978). Why astrology is a pseudoscience. *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1978, 223-234. <http://www.jstor.org/stable/192639>
- Thelwall, M. (2021). Measuring societal impacts of research with altmetrics? Common problems and mistakes. *Journal of Economic Surveys*, 35(5), 1302-1314. <https://doi.org/10.1111/joes.12381>
- Thelwall, M., & Fairclough, R. (2017). The research production of nations and departments: A statistical model for the share of publications. *Journal of Informetrics*, 11(4), 1142-1157. <https://doi.org/10.1016/j.joi.2017.10.001>
- Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, C. R. (2013). Do altmetrics work? Twitter and ten other social web services. *PLoS One*, 8(5), e64841. <https://doi.org/10.1371/journal.pone.0064841>
- Tillett, B. B. (2001). Bibliographic relationships. In C. A. Bean, & R. Green (Eds.), *Relationships in the organization of knowledge* (pp. 19-35). Springer.
- Tokar, A. (2012). Semantic change of the publication-concept? In A. Tokar, M. Beurskens, S. Keuneke, M. Mahrt, I. Peters, C. Puschmann, T. van Treeck, & K. Weller (Eds.), *Science and the Internet* (pp. 193-205). Düsseldorf University Press.
- Toutkoushian, R. K., Porter, S. R., Danielson, C., & Hollis, P. R. (2003). Using publications counts to measure an institution's research productivity. *Research in Higher Education*, 44(2), 121-148. <https://doi.org/10.1023/A:1022070227966>
- Tüür-Fröhlich, T. (2016). *The non-trivial effects of trivial errors in scientific communication and evaluation*. Verlag Werner Hülsbusch.
- Van Noorden, R. (2010). Metrics: A profusion of measures. *Nature*, 465(7300), 864-866. <https://doi.org/10.1038/465864a>
- van Raan, A. (2005). Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics*, 62(1), 133-143. <https://doi.org/10.1007/s11192-005-0008-6>
- Vera-Baceta, M. A., Thelwall, M., & Kousha, K. (2019). Web of Science and Scopus language coverage. *Scientometrics*, 121(3), 1803-1813. <https://doi.org/10.1007/s11192-019-03264-z>
- Verleysen, F. T., & Engels, T. C. E. (2018). How arbitrary are the weights assigned to books in performance-based research funding? An empirical assessment of the weight and size of monographs in Flanders. *ASLIB Journal of Information Management*, 70(6), 660-672. <https://doi.org/10.1108/AJIM-05-2018-0110>
- Vílchez-Román, C. (2017). BRICS' scientific excellence and the search for relevance and replicability. *Journal of the Association for Information Science and Technology*, 68(2), 539. <https://doi.org/10.1002/asi.23654>
- Waltman, L., & van Eck, N. J. (2019). Field normalization of scientometric indicators. In W. Glänzel, H. F. Moed, U. Schmoch, & M. Thelwall (Eds.), *Springer handbook of science and technology indicators* (pp. 281-300). Springer.
- Waltman, L., van Eck, N. J., Visser, M., & Wouters, P. (2016). The elephant in the room: The problem of quantify-

- ing productivity in evaluative scientometrics. *Journal of Informetrics*, 10(2), 671-674. <https://doi.org/10.1016/j.joi.2015.12.008>
- Wang, K., Shen, Z., Huang, C., Wu, C. H., Dong, Y., & Kanakia, A. (2020). Microsoft academic graph: When experts are not enough. *Quantitative Science Studies*, 1(1), 396-413. https://doi.org/10.1162/qss_a_00021
- Watson, C. (2022). Rise of the preprint: How rapid data sharing during COVID-19 has changed science forever. *Nature Medicine*, 28(1), 2-5. <https://doi.org/10.1038/s41591-021-01654-6>
- Weingart, P., & Winterhager, M. (1984). *Die Vermessung der Forschung. Theorie und Praxis der Wissenschaftsindikatoren*. Campus. German.
- Zahedi, Z., & Costas, R. (2018). General discussion of data quality challenges in social media metrics: Extensive comparison of four major altmetric data aggregators. *PLoS One*, 13(5), e0197326. <https://doi.org/10.1371/journal.pone.0197326>
- Zhao, D., Cappello, A., & Johnston, L. (2017). Functions of uni- and multi-citations: Implications for weighted citation analysis. *Journal of Data and Information Science*, 2(1), 51-69. <https://doi.org/10.1515/jdis-2017-0003>
- Zimmer, F., Scheibe, K., & Stock, W. G. (2018, July 15-20). A model for information behavior research on social live streaming services (SLSSs). In G. Meiselwitz (Ed.), *Proceedings of the 10th International Conference, SCSM 2018: Social Computing and Social Media. Technologies and Analytics* (pp. 429-448). Springer.
- Zitt, M., Ramanana-Rahary, S., & Bassecoulard, E. (2005). Relativity of citation performance and excellence measures: From cross-field to cross-scale effects of field-normalisation. *Scientometrics*, 63(2), 373-401. <https://doi.org/10.1007/s11192-005-0218-y>