On the Value and Use of Metrics and Rankings: a Position Paper

© Yannis Manolopoulos
Department of Informatics, Aristotle University
54124 Thessaloniki, Greece
manolopo@csd.auth.gr

Abstract

In this paper we provide an introduction to the field of Bibliometrics. In particular, first we briefly describe its beginning and its evolution; we mention the main research fora as well. Further we categorize metrics according to their entity scope: metrics for journals, conferences and authors. Several rankings have appeared based on such metrics. It is argued that these metrics and rankings should be treated with caution, in a light relative way and not in an absolute manner. Primarily, it is the human expertise that can rigorously evaluate the above entities.

1 Introduction

The term "*Bibliometrics*" has been proposed by Alan Pritchard in 1969 [38]. According to Wikipedia, "Bibliometrics is statistical analysis of written publications, such as books or articles" [7]. A relevant field is "*Scientometrics*", a term coined by Vasily Nalimov in 1969 [36]. According to Wikipedia, "Scientometrics is the study of measuring and analysing science, technology and innovation" [41]. Finally, "*Citation analysis*" is a fundamental tool for Bibliometrics and deals with the "examination of the frequency, patterns, and graphs of citations in documents" [12].

A milestone in the development of the field of Bibliometrics was the introduction of the "*Journal Impact Factor*" (IF) by Eugene Garfield in 1955 [22]. Garfield founded the Institute for Scientific Information (ISI) in 1964, which published the Science Citation Index and the Social Science Citation Index. ISI was acquired by Thomson Reuters in 1992.

For about four decades IF was the standard tool for academic evaluations. Despite the fact that it was proposed as a metric to evaluate journals' impact, it was used as a criterion to evaluate the quality of scholarly work by academicians and researchers as well. It was only in 2005 that Jorge Hirsh, a physicist, proposed the *h-index* as a simple and single number to evaluate the production and the impact of a researcher's work [29]. During the last 10-15 years the field has flourished and significant research has appeared in competitive journals and conferences.

Based on these metrics, several rankings have appeared in the web, e.g. for journals, conferences and authors. On the other hand, university rankings appear in popular newspapers; actually, they are the beloved topic of journalists and politicians. University rankings are commercial artifacts of little scientific merit as they are based on arbitrary metrics (like the ones previously mentioned) and on other unjustified subjective criteria.

The purpose of this position paper is to explain that these metrics and rankings should be used with great skepticism. To a great extent, they shed light only to some particular facets of the entity in question (be it a journal, an author etc.); moreover, they are often contradictory to each other. The suggestion is to use this information with caution and pass it through an expert's filtration to come up with a scientific and objective evaluation.

The rest of the paper has the following structure. In the next section we provide more information about the Bibliometric/Scientometric community. Then, we introduce and annotate several metric notions for the evaluation of journals, conferences and authors. In Section 4 we mention assorted rankings and pinpoint their contradictions, limitations and fallacies. We devote Section 5 to discussing university rankings. In the last section we introduce the Leiden manifesto, an article that tries to put academic evaluations in an academic (not commercial, not mechanistic) framework, and in conclusion we state the morals of our analysis. Due to space limitations, this position paper is far from an exhaustive study on the subjects mentioned.

2 The Bibliometrics Community

During the last two decades a new research community was formed focusing on bibliometric and scientometric issues.

The research output of this community appears in specialized journals and conferences. In particular we note the following journals: (i) Journal of the Association for Information Science and Technology by Wiley, (ii) Scientometrics by Springer, (iii) Journal of Informetrics by Elsevier, (iv) Journal of Scientometrics and Information Management by Taylor and Francis, (v) Research Evaluation by Oxford Journals.

Two major annual conferences are running for more than a decade. For example, the International Conference of the International Society for Scientometrics and Informetrics (ISSI) is organizing its 16th event at Yuhan/China in October 2017, whereas the 12th International Conference on Webometrics, Informetrics, and Scientometrics (WIS) of the COLLNET community is

organized in December 2016 at Nancy/ France. In addition, assorted papers appear in other major outlets related to Artificial Intelligence, Data Mining, Information Retrieval, Software and Systems, Web Information Systems, etc.

Notably, there are several major databases with bibliographic data. Among others we mention: Google Scholar [25] and the tool Publish or Perish [39], which runs on top of Google Scholar, MAS (Microsoft Academic Search) [35], Scopus [43] by Elsevier and Web of Science [53] (previous known as ISI Web of Knowledge) by Thomson Reuters and DBLP (Data Bases and Logic Programming) [17] by the University of Trier.

3 The Spectrum of Metrics

3.1 Impact Factor

As mentioned earlier, IF is the first proposed metric aiming at evaluating the impact of journals. For a particular year and a particular journal, its IF is the average number of citations calculated for all the papers that appeared in this journal during the previous 2 years. More specifically, this is the 2-years IF as opposed to the 5-years IF, which has been proposed relatively recently as a more stable variant.

IF is a very simple and easily understood notion; it created a business, motivated researchers and publishers and was useful for academicians and librarians. However, IF has been criticized for several deficiencies. For instance,
o it is based mostly on journals in English,
o it considers only a fraction of the huge set of peer-reviewed journals,
o it did not take under account (until recently) conference proceedings, which play an important role in scholar communication in computer science for example,
o it fails to compare journals across disciplines,
o it is controlled by a private institution and not by a democratically formed scientific committee, etc.
On top of these remarks, studies suggest that citations are not *clean* and therefore the whole structure is weak [34]. Also, a recent book illustrates a huge number of flaws encountered in IF measurements [49].

IF can be easily manipulated by the journal's editors-in-chief, who are under pressure in the competitive journal market. For example, the editors-in-chief:
o may ask from authors to add extra references of the same journal,
o may invite/accept surveys as these articles attract more citations that regular papers, or
o may prefer to publish articles that seem to be well cited in the future.

There is yet another very strong voice against the over-estimation of IF. Philip Campbell, Editor-in-Chief of the prestigious *Nature* journal, discovered that few papers make the difference and increase the IF of a specific journal. For example, the IF value of Nature for the year 2004 was 32.2. When Campbell analyzed Nature papers over the relevant period (i.e., 2002-2003), he found that 89% of the impact factor was generated by just 25% of the papers [11]. This is yet another argument against the use of IF to evaluate authors. That is, an author should not be proud just because he published an article in a journal with high IF; on the contrary he should be proud if his paper indeed contributed in this high IF value. However, it is well known that the distribution of the number of citations per paper is exponential [26]; therefore, most probably the number of citations of a paper per year will be less than half of the IF value.

In this category another score can be assigned: the *Eigenfactor* developed by Jevin West and Carl Bergstrom of the University of Washington [6]. As mentioned in Wikipedia: "The Eigenfactor score is influenced by the size of the journal, so that the score doubles when the journal doubles in size" [21]. Also, Eigenfactor score has been extended to evaluate the impact at the author's level.

More sophisticated metrics have been proposed, not only for reasons of elegance but also in the course of commercial competition as well.

3.2 Metrics by Elsevier Scopus

It is well-known that IF values range significantly from one field to another. There are differences in citation practices, in the lag time between publication and its future citation and in the particular focus of digital libraries. It has been reported that "the field of Mathematics has a weighted impact factor of IF=0.56, whereas Molecular and Cell Biology has a weighted impact factor of 4.76 - an eight-fold difference" [4].

Scopus, the bibliometric research branch of Elsevier uses two important new metrics: *SNIP* (Source Normalized Impact per Paper) [47] and *SJR* (Scimago Journal Rank) [42]. Both take into account two important parameters.

SNIP has been proposed by the Leiden University Centre for Science and Technology Studies (CWTS) based on the Scopus database. According to SNIP citations are normalized by field to eliminate variations; IFs are high in certain fields and low in others. SNIP is a much more reliable indicator than the IF for comparing journals among disciplines. It is also less open to manipulation. Therefore, normalization is applied to put things in a relative framework and facilitate the comparison of journals of different fields.

On the other hand, SJR has been proposed by the SCImago research group from Consejo SCImago research group from the Consejo Superior de Investigaciones Científicas (CSIC), University of Granada, Extremadura, Carlos III (Madrid) and Alcalá de Henares. SJR indicates which journals are more likely to have articles cited by prestigious journals, not simply which journals are cited the most, adopting a reasoning similar to that of Pagerank's algorithm [10].

These metrics have been put into practice as the reader can verify by visiting websites of journals published by Elsevier.

3.3 Metrics for Conferences

Conferences are not treated in a uniform way from one discipline to another and even from subfield to subfield. For example, there are conferences where only abstracts are submitted, accepted and published in a booklet, whereas there are other conferences where full papers are submitted and reviewed by ~3 referees in the same manner as journals. Apparently, the latter articles can be treated as first class publications, in particular if the acceptance ratio is as high as 1 out 5, or 1 out of 10 as it happens in several prestigious conferences (such as WWW, SIGMOD, SIGKDD, VLDB, IJCAI, etc.).

Thus, an easy metric to evaluate the quality of a conference is the acceptance ratio (i.e. number of accepted vs. number of submitted papers). Several publishing houses (e.g. Springer) specify that the acceptance ratio should be >33%. It is a common practice to report such numbers in the foreword of conference proceedings.

Several websites collect data about the acceptance ratios of conferences of several fields such as Theoretical CS, Computer Networks, Software Engineering etc. [1-3]. In general, there is a trend towards events with stricter acceptance policies. The humoristic study of [16] seriously deconstructs such approaches.

Apart from the acceptance ratio, an effort to quantify the impact of conferences has been first initiated by Citeseer, a website and service co-created by Lee Giles, Steve Lawrence and Kurt Bollacker at NEC Research Institute [23]. In particular, using its own datasets Citeseer calculates the IF of a rich set journals and conferences in a unique list. Nowadays, Citeseer is partially maintained by Lee Giles at the Pennsylvania State University [13]; practically, it has been surpassed by Google Scholar.

3.4 Metrics for Authors

In 1985, Jorge Hirsch, a physicist at UCSD, invented the notion of the *h-index* [29]. According to Wikipedia: "a scholar with an index of *h* has published *h* papers each of which has been cited in other papers at least *h* times" [30]. Thus, the *h*-index illustrates both the production and the impact of a researcher's work. It is not just a single number but a 2-dimensional number. *h*-index was a breakthrough; it was a brand new notion that broke the monopoly of IF in academic evaluations.

A propos, it came up that the *h*-index was just a re-invention of a similar metric. Arthur Eddington, an English astronomer, physicist, and mathematician of the early 20th century, was an enthusiastic bicycler. In the context of cycling, Eddington's number is the maximum number *E* such that the cyclist has cycled *E* miles on *E* days. Eddington's own *E*-number was 84 [20].

Although a breakthrough notion, the *h*-index received some criticism when it was put in practice. In particular, the following issues were brought up:
o it does not consider peculiarities of each specific field,
o it does not consider the order of an author in the list of authors,
o it has a reduced discriminative power as it is an integer number,
o it can be manipulated with self-citations, which cannot be revealed in Google Scholar,
o it has a correlation with the number of the author's publications,
o it constantly increases with time and cannot show the progress or stagnation of an author.

Soon after the invention of the h-index, the field of Bibliometrics flourished and a lot of variants were proposed. The following is only a partial list: g-index, a-index, h(2)-index, hg-index, q^2-index, r-index, ar-index, m-quotient, k-index, f-index, m-index, h_w-index, hm-index, h_{rat}-index, v-index, e-index, π-index, RC-index, CC-index, ch-index, n-index, p-index, w-index, and so on and so forth [31]. The present author's team proposed the following 3 variants: contemporary h-index, trend h-index, normalized h-index [45]. After this flood of variants, several such studies were reported aiming at analysing, comparing and categorizing the multiplicity of indicators [8,9,54,55].

In passing, there have been efforts in studying and devising metrics for the whole citation curve of the works by an author, as a metric supplementary to the *h*-index. In this direction, we proposed two new metrics: the perfectionism index [46] and the fractal dimension [24] to penalize long tails and to dissuade authors from writing papers of low value.

The books by Nikolay Vitanov [52] and Roberto Todeschini, Alberto Baccini [48] give extensive insight into these metrics for authors. In addition, in Publish or Perish [39], a website (maintained by Harzing) and book [27], the most common of these variants have been implemented. Finally, Matlab includes several such implementations as well.

4 The Spectrum of Rankings

Ranking is a popular game in academic environments. One can easily find rankings about authors, journals, conferences, and universities as well. Here, we comment on some interesting rankings drawn from several websites. In particular, we will comment on university rankings in the next section.

DBLP website [17] is maintained by Michael Ley at the University of Trier. As of May 2016, its dataset contains more than 1.7 million authors and 3.5 million articles. Based on this dataset, DBLP posts a list of prolific authors in terms of publications of all sorts, e.g. journal and conference papers, books, chapters in books etc. It is interesting to note that Vincent Poor, a researcher at Princeton University, is the most productive person in this ranking with an outcome of 1348 publication (as of 21/10/2016) [18]. Another ranking with the same dataset ranks authors according to the average production per year. Vincent Poor can be found in the 19th position in this ranking (as of 21/10/2016) [19].

Jens Palsberg of UCLA maintains a website where, by using the DBLP datasets, a list of authors ordered according to decreasing *h*-index is produced. First name in this list is Herbert Simon, a Professor at CMU, Nobel

Laureate, Turing Award recipient, ACM Fellow; his *h*-index is 164. In this list, Vincent Poor appears with *h*-index equal to 70 [32].

MAS provides a variety of rankings using a dataset of 80 million of articles [35]. For example, it provides two ranked lists of authors according to productivity and according to impact. When the whole dataset is taken into account, then in terms of the number of publications Scott Shenker is 1st, Ian Foster is 2nd and Hector Garcia-Molina is 3rd. According to the number of citations Ian Foster is 1st, Ronald Rivest is 2nd and Scott Shenker is 3rd. Other ranking can be produced by limiting the time window during the last 5 or 10 years. For instance, Vincent Poor is ranked 2nd in terms of productivity for the period of the last 10 years.

Similarly, MAS provides rankings of conferences according to the number of publications or citations for certain periods (i.e. 5 years, 10 years, or the whole dataset). Steadily, INFOCOM, SIGGRAPH, CVPR, ICRA, ICASSP appear at the top.

In an analogous manner, MAS provides rankings for journals. When considering the whole data set, the top journals are CACM, PAMI and TIT. During the last 5 years, new fields came up and, thus, new journals gained acceptance: see for example Expert Systems with Applications and Applied Soft Computing. It is important to notice that these rankings use raw numbers, i.e. without any normalization. However, they show trends in science with time.

The above paragraphs show that there are several kinds of ranking, each with a different emphasis and as such they should be treated with caution.

Another example of misuse of rankings concerns the classification of journals and conferences. CORE is an Australian website/service, where journals and conferences are divided in 5 categories as illustrated in the following table [15]. Numbers show the percentages of journals or conferences at their *corresponding* category. Similar categorizations exist in other websites. Even though it is not transparent how the percentages were calculated and the rankings are based on somewhat arbitrary listings and categorizations, such rankings have great acceptance and in several instances state funding may be based on them.

	A*	A	B	C	Other
Journals	7%	17%	27%	46%	3%
Conferences	4%	14%	26%	51%	5%

We give another example where caution is needed. We present two tables. The first table contains data from Aminer [5], which runs on top of DBLP. This table shows the top-10 outlets for "database and data mining" sorted by the *h5*-index, a variation of *h*-index for journals. *H5* is the largest number h such that h articles published in 2011-2015 have at least h citations each.

In an analogous manner, the following table gives the top-10 outlets for "Database and Data Mining" sorted by *h5* according to Google Scholar for "Database and Information System".

1	WWW conference	66
2	Information Sciences	62
3	ACM KDDM	56
4	IEEE TKDE	53
5	ACM WSDM conference	50
6	JASIST	47
7	ACM SIGIR	42
8	IEEE ICDE conference	40
9	ACM CIKM conference	38
10	IEEE ICDM conference	33

1	WWW conference	74
2	VLDB conference	67
3	IEEE TKDE	66
4	arXiv Social & Infor. Networks (cs. SI)	66
5	ACM SIGMOD conference	65
6	arXiv Databases (cs DB)	61
7	ICWSM (weblog) conference	60
8	ICWSM (web) conference	59
9	ACM WSDM conference	58
10	IEEE ICDE conference	52

We note that the two lists have only 5 items in common, and in different order. At first, one might think that the two lists were not comparable because since they were produced by querying different keywords. However, since the first table contains outlets related to Information Systems, whereas the second one contains outlets related to Data Mining the two lists are indeed comparable. This example illustrated that the adoption of a ranking versus another is a subjective matter.

5   University Rankings

Nowadays education is considered as a product/service and, thus, there is a growing financial interest in this global market. Universities try to improve their position in the world arena. Thus, university rankings try to satisfy the need of universities for visibility. These ranking are a popular topic for journalists and, therefore, for politicians as well. However, beforehand we claim that there is little scientific merit in these rankings.

Some rankings are widely-known from mass media. We mention alphabetically the most commercial ones:
o Academic Ranking of World Universities (Shanghai),
o QS World University Rankings (QS),
o Times Higher Education World University Rankings (THE).

Other rankings originate from academic research teams, such as:
o Leiden Ranking,
o Professional Ranking of World Universities (École Nationale Supérieure des Mines de Paris),
o SCImago Institutions Ranking,
o University Ranking by Academic Performance (Middle East Technical University),
o Webometrics (Spanish National Research Council),
o Wuhan University.

A full list of such rankings exists at Wikipedia [50].

University rankings are intensively criticized for a number of reasons. For example:

o All rankings are based on a number of subjective criteria.
o In all cases, the choice of each particular criterion and its weight are arbitrary.
o To a great extent, these criteria are correlated.
o Evaluation for some criteria is based on surveys, e.g. "academic reputation" or "employer reputation" by QS, which count for 50% of the total weight. The same holds for the "reputation survey" by THE, which counts for 17.9% or 19.5% or 25.3%, if the examined institution is a medical, an engineering or an arts/humanities school, respectively. Such surveys are totally not-transparent.
o THE devotes a 7.5% of the total weight for the international outlook, sub-categorized into "ratio of international to domestic staff", "international co-authorship" and "ratio of international to domestic students". In the same way, QS considers "international student ratio" and "international staff ratio" with a special weight of 5%+5%. Clearly, such criteria favour Anglo-Saxon universities.
o The number of publications and the number of citations (without normalization) favour big universities; this is probably a reason for a general trend in merging universities in Europe.
o No ranking considers whether a university is an old or a new institution, big or small, a technical university or a liberal arts one, etc. Thus, different entities are compared.
o In general, ranking results are not reproducible, an absolutely necessary condition to accept an evaluation as methodologically reliable.
o QS adopts the h-index at a higher level, i.e. not at the author's level but for a group of academicians. This is beyond the fair use of the original idea by Hirsch since it does not consider the size of the examined institution, neither it performs any normalization.
o The rankings exert influence on researchers to submit papers to "prestigious" journals (e.g. Nature, Science). Since such journals follow particular policies as to what is in fashion researchers may not work on what they truly think is worthwhile but according to external/political criteria acting as sirens [44].
o Finally, and probably the most important point of this criticism is that university rankings are misleading proportionally to the degree that they are based on (a) collections of citations from English-language digital libraries, (b) erroneous collections of citations, (c) IF calculations, which ignore whole statistical distributions of a single number, (d) higher level h-index calculations, which are conceptually wrong. In other words, "garbage in, garbage out".

All rankings are not equally unacceptable. Several independent studies agree that ARWU is probably the most reliable in comparison to other commercial rankings [33], whereas QS is the most criticized ranking. On the other hand, between the rankings originated from academic institutions, Leiden is considered as the most reliable as it stems from a strong research team with significant academic reputation and tradition in the field of Bibliometric/Scientometrics. On the other hand, the ranking of Webometrics is criticized for the adoption of non-academic criteria, such as the number of web pages and files and their visibility and impact according to the number of received inlinks.

Based on the above discussion, one can understand why the question "science or quackery" arises [40]. In a recent note by Moshe Vardi, Editor-in-Chief of CACM and professor with Rice University, same scepticism was reported [51]. Moreover, some state authorities are critical against these methodologies [37]. However, it is sad that rankings are "*here to stay*" because strong financial interests worldwide support such approaches.

At this point, we mention a very useful website which, based on the DBLP dataset, ranks American CS departments in terms of the number of faculty and the number of publications in selected fora, by picking certain CS subfields [14].

6 Discussion and Morals

The intention of this position paper is the following. Bibliometrics is a scientific field supported by a strong research community. Although the term is not new, during the last years there is an intense research in the area due to the web and open/linked data.

The outcome of Bibliometrics is most often misused by mass media and journalists, state authorities and politicians, and even in the academic world. Criticism has been expressed for several metrics and rankings, not without a reason.

In 2015, a paper was published in Nature: "the Leiden Manifesto for research metrics". More specifically, the paper states 10 principles to guide research evaluation [28]. We repeat them here in a condensed style:
1. Quantitative evaluation should support qualitative, expert assessment.
2. Measure performance against the research missions of the institution, group or researcher.
3. Protect excellence in locally relevant research.
4. Keep data collection and analytical processes open, transparent and simple.
5. Allow those evaluated to verify data and analysis.
6. Account for variation by field in publication and citation practices.
7. Base assessment of individual researchers on a qualitative judgement of their portfolio.
8. Avoid misplaced concreteness and false precision.
9. Recognize the systemic effects of assessment and indicators.
10. Scrutinize indicators regularly and update them.

Probably, the last principle is the most important. Since it is easy for humans to cleverly adapt to external rules and to try to get the most benefit out of them, the Bibliometrics community has to devise and promote new metrics for adoption by academia and others.

Finally, we close this paper with a proposed list of do's and don'ts.

1. Do not evaluate researchers based on the number of publications or the IF of the journals they appeared.
2. Evaluate researchers with their *h*-index and variants (resolution according to competition).
3. To further evaluate researchers, focus on the whole citation curve and its tail in particular (relevant metrics: perfectionism index and fractal dimension).
4. Do not evaluate journals based on their IF
5. Evaluate journals with the SCIMAGO and EIGEN-FACTOR scores as they are robust and normalized.
6. Further, ignore journal metrics and choose to work on the topics that inspire you.
7. Metrics are not panaceas; metrics should change periodically.
8. Do not get obsessed with contradictory rankings for authors, journals and conferences.
9. Ignore university rankings; they are non-scientific, non-repeatable, commercial, unreliable.
10. Follow your heart and research what attracts and stimulates you.

Acknowledgments

References

1. Acceptance ratio of TCS conferences http://www.lamsade.dauphine.fr/~sikora/ratio/confs.php
2. Acceptance ratio of Networking conferences https://www.cs.ucsb.edu/~almeroth/conf/stats/
3. Acceptance ration of SW Engineering conferences http://taoxie.cs.illinois.edu/seconferences.htm
4. Althouse B., West J., Bergstrom T. and Bergstrom C.: "Differences in Impact Factor across Fields and Over Time", *Journal of the American Society on Information Sciences & Technology*, Vol.60, No.1, pp.27-34, 2009.
5. Aminer https://aminer.org/ranks/conf
6. Bergstrom C.T., West J.D. and Wiseman M.A.: "The Eigenfactor Metrics", *Journal of Neuroscience*, Vol.28, No.45, pp.11433–11434, 2008.
7. Bibliometrics: https://en.wikipedia.org/wiki/Bibliometrics
8. Bollen J., van de Sompel H., Hagberg A., and Chute R.: "A Principal Component Analysis of 39 Scientific Impact Measures", *PLOS One*, 4, e6022, 2009.
9. Bornmann L. et al.: "A Multilevel Meta-analysis of Studies Reporting Correlations Between the *h*-index and 37 Different *h*-index Variants", *Journal of Informetrics*, Vol.5, No.3, pp.346-359, 2011.
10. Brin S. and Page L.: "The Anatomy of a Large-scale Hypertextual Web Search Engine", *Computer Networks and ISDN Systems*, Vol.30, No.1-7, pp.107-117, 1998.
11. Campbell P.: "Escape from the Impact Factor", *Ethics in Science and Environmental Politics*, Vol.8, pp.5-7, 2008.
12. Citation Analysis: https://en.wikipedia.org/wiki/Citation_analysis
13. CiteSeer Digial Library http://citeseerx.ist.psu.edu/index
14. Computer Science ranking http://csrankings.org/
15. Computing Research and Evaluation (CORE) http://www.core.edu.au/
16. Cormode G., Czumaj A. and Muthukrishnan S.: "How to Increase the Acceptance Ratios of Top Conferences?", http://www.cs.rutgers.edu/~muthu/ccmfun.pdf
17. DBLP http://dblp.uni-trier.de/
18. DBLP, prolific authors http://dblp.uni-trier.de/statistics/prolific1
19. DBLP, prolific authors per year http://dblp.l3s.de/browse.php?browse=mostProlificAuthorsPerYear
20. Eddington Arthur https://en.wikipedia.org/wiki/Arthur_Eddington
21. Eigenfactor metric https://en.wikipedia.org/wiki/Eigenfactor
22. Garfield E.: "Citation Indexes for Science: a New Dimension in Documentation through Association of Ideas", *Science*, 122, 108-111, 1955.
23. Giles C.L., Bollacker K. and Lawrence S.: "Cite-Seer: An Automatic Citation Indexing System", *Proceedings 3rd ACM Conference on Digital Libraries*, pp.89-98, 1998.
24. Gogoglou A., Sidiropoulos A., Katsaros D. and Manolopoulos Y.: "Quantifying an Individual's Scientific Output Using the Fractal Dimension of the Whole Citation Curve", *Proceedings 12th International Conference on Webometrics, Informetrics & Scientometrics (WIS),* Nancy, France, 2016.
25. Google Scholar www.scholar.google.com
26. Gupta H., Campanha J. and Pesce R.: "Power-law Distributions for the Citation Index of Scientific Publications and Scientists", *Brazilian Journal of Physics*, Vol.35, No.4a, pp.981-986, 2005.
27. Harzing A.W.: "*Publish or Perish*", Tarma Software Research, 2010.
28. Hicks D., Wouters P., Waltman L., de Rijke S. and Rafols I.: "The Leiden Manifesto for Research Metrics", Nature, April 2015
29. Hirsch J. E. "An Index to Quantify an Individual's Scientific Research Output", *Proceedings National Academy of Sciences*, Vol.102, No.46, pp.16569–16572, 2005.
30. *h*-index https://en.wikipedia.org/wiki/H-index

31. *h*-index variants
    http://sci2s.ugr.es/hindex
32. *h*-index for CS scientists
    http://web.cs.ucla.edu/~palsberg/h-number.html
33. Lages J., Patt A. and Shepelyansky D.: "Wikipedia Ranking of World Universities", *Arxiv*, 2016.
34. Lee D., Kang J., Mitra P., Giles L. and On B.W.: "Are Your Citations Clean?", *Communications of the ACM*, Vol.50, No.12, pp.33-38, 2007.
35. Microsoft Academic Search
    http://academic.research.microsoft.com/
36. Nalimov V., Mul'chenko Z.M.: "Наукометрия, Изучение развития науки как информацио-нного процесса" [Naukometriya, the study of the development of science as an information process], *Nauka*. p.191, Moscow, 1969.
37. Norwegian universities
    http://www.universityworldnews.com/article.php?story=20140918170926438
38. Pritchard A.: "Statistical Bibliography or Bibliometrics?", *Journal of Documentation*, Vol.25, No.4, p.348-349, 1969.
39. Publish or Perish
    http://www.harzing.com/pop.htm
40. Science or quackery
    https://www.aspeninstitute.it/aspenia-online/article/international-university-rankings-science-or-quackery
41. Scientometrics
    https://en.wikipedia.org/wiki/Scientometrics
42. SCIMAGO
    http://www.scimagojr.com/
43. Scopus
    www.scopus.com
44. Schekman R.
    https://www.theguardian.com/science/2013/dec/09/nobel-winner-boycott-science-journals
45. Sidiropoulos A., Katsaros D. and Manolopoulos Y.: "Generalized Hirsch *h*-index for Disclosing Latent Facts in Citation Networks, *Scientometrics*, Vol.72, No.2, pp.253-280, 2007.
46. Sidiropoulos A., Katsaros D. and Manolopoulos Y.: "Ranking and Identifying Influential Scientists vs. Mass Producers by the Perfectionism Index", *Scientometrics*, Vol.103, No.1, pp.1-31, 2015.
47. SNIP
    http://www.journalindicators.com/
48. Todeschini R. and Baccini A.: "*Handbook of Bibliometric Indicators*", Wiley, 2016.
49. Tüür-Fröhlich T.: "*The Non-trivial Effects of Trivial Errors in Scientif Communication and Evaluation*", Verlag Werner Hülsbusch, 2016.
50. University rankings
    https://en.wikipedia.org/wiki/College_and_university_rankings
51. Vardi M.: "Academic Rankings Considered Harmful!", *Communications of the ACM*, Vol.59, No.9, p.5, 2016
52. Vitanov N.: "*Science Dynamics and Research Production*", Springer, 2016.
53. Web of Science
    http://ipscience.thomsonreuters.com/
54. Wildgaard L., Schneider J.W. and Larsen B.: "A Review of the Characteristics of 108 Author-level Bibliometric Indicators", *Scientometrics*, Vol.101, pp.125-158, 2014.
55. Yan Z., Wu Q. and Li X.: "Do Hirsch-type Indices Behave the Same in Assessing Single Publications? An Empirical Study of 29 Bibliometric Indicators", *Scientometrics*, 2016.