Handbook Bibliometrics

Handbook Bibliometrics

Edited by
Rafael Ball

DE GRUYTER
SAUR

ISBN 978-3-11-064227-8
e-ISBN (PDF) 978-3-11-064661-0
e-ISBN (EPUB) 978-3-11-064259-9

Library of Congress Control Number: 2020945307

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

Cover image: imaginima / E+ / Getty Images
Printing and binding: CPI books GmbH, Leck

www.degruyter.com

Contents

Introduction

Rafael Ball

The development of academia, its structure and systems, its communication, and above all its questions to the world is a fascinating research field in itself. The fact that, over its 2,000-year history, academia has blossomed from an individual commitment and the interests of individuals into an institutionalised form is of particular importance. The foundation of academies and universities at the end of the Middles Ages, the enthusiasm for new things and the open-mindedness during the Renaissance, still fascinate to this day. Universities and all the other parallel structures of institutionalised academia and research evolved into a veritable "mass phenomenon" in the early twentieth century. A growing number of people are interested in the vast range of issues concerning humans, society and the world around them. However, ever more people also have the option of choosing academia for a living and research as a career, which is yielding ever more answers, insights and new products. Research and academic questions are going into increasing depth. However, the horizon in which the findings need to be embedded is also becoming broader and more complex; for many issues, only collaboration between the various disciplines produces the right solutions. A network of comprehensive answers to academic questions is taking shape, although the academic knowledge process will never be completed. Academia does not provide definitive answers; they are only ever valid until new findings are able to refute, modify or improve on the explanation. Karl Popper describes this as the falsification theory in his book *The Logic of Scientific Research* (first published as *Logik der Forschung* in 1934; English translation 1959).[1]

According to Popper, one fundamental principle of academia is that it shares its results, discusses with other scientists and thus keeps moving forward. However, this also means that academic results are never generated purely for one's own personal development, but always have to be published as well. Otherwise, in the best case scenario it involves research in private industry to develop products and, in the worst case, "secret science", exactly like the "maverick scholars" portrayed in poor television series.

Since the mid-twentieth century, the number of publications has boomed. Increasing numbers of scientists are producing more and more findings and publishing more and more results in books and journals, at conferences and, ever since the early 2000s, in or on blog entries, websites, databases or in other state of the art electronic forms of (social) media, such as the meanwhile highly successful social academic networks that are increasingly being used. The volume of academic output has

1 Karl Popper, Logik der Forschung, Akademieverlag, 2013.

Dr. **Rafael Ball**, Director ETH Libraries and Collections, ETH Zurich, Rämistrasse 101, 8092 Zuerich, Switzerland. Email: rafael.ball@library.ethz.ch.

https://doi.org/10.1515/9783110646610-002

reached a scale which the individual often can no longer fathom. Already today, twenty per cent of all academic publications in Germany are not even cited once. The number of journals, books and conferences has ballooned. Digital publishing formats and social media are being discovered as additional or new publishing platforms and used massively.

On the one hand, this has created a wonderful situation for knowledge gains in academia; on the other hand, however, it has become impossible to obtain an overview (even a rough one) of the results published and their quality and importance for the respective subject.

In this respect, the bell tolls for bibliometrics. Anyone who is no longer able to read or perceive everything that is published in or even beyond his or her discipline has to choose. Due to the sheer volume, however, in most cases a quantitative selection is no longer possible. Bibliometrics offers a way out here. It gauges the perception of academic publications via the roundabout route that citations take and uses this to draw inferences about their quality. If academic publications are cited frequently, it is safe to assume that they are important papers (and worth reading in detail). By the same token, a paper that is cited seldomly, if at all, seems less relevant.

This fundamental principle of inferring the quality of a paper indirectly from the number of citations is the basis of bibliometrics, the requirements, methods and results of which are the subject of this handbook.

The first attempts to establish a link between citations of publications and the quality of the papers were made as early as 1927.[2] However, the actual development and eventual broad usage of bibliometrics only really took off with the advent of the digital age. Only in this way can a mass evaluation of analysis data on citations made be realised. The fact that this spawns a broad field of statistical "variant forms" meant it was inevitable that bibliometrics itself would blossom into a (virtually incomprehensible) science. Nowadays, bibliometrics refers to a vast number of forms of metrics, which measure the perception and importance of academic publications in statistical form. In the process, today citations are merely one gauge among many other attention indicators. As a result, however, the consideration of bibliometrics has long since become part of scholarly communication and must also be embedded in current developments such as open access or open science.

This handbook examines all these different topics related to bibliometrics.

In the chapter entitled 'History and Institutionalization of Bibliometrics', articles on the historical development of bibliometrics and its methods are described and explained. The institutionalization of bibliometrics in the course of this historical development also features extensively, based on selected people and institutions that can be regarded as milestones of bibliometrics. They have been instrumental in bib-

2 Gross, PLK & Gross, EM 1927, College libraries and chemical education, *Science* 66, pp. 385–389.

liometrics becoming an accepted part of the variety of methods used to evaluate academia.

The chapter 'Theory, Principles and Methods of Bibliometrics' addresses the methodical foundations of bibliometrics, which are essentially mathematical and statistical in character. An attempt is also made at a heuristic assessment of what measuring attention can actually tell us about the quality of academic publications (and what it cannot).

Besides offering an overview of the basic indicators and their importance and calculation, the chapter '(Classical) Indicators' also singles out individual core indicators such as the h-Index and analyses them in depth.

The chapter 'Alternative Metrics (Altmetrics)' focuses on social media and their metrics, which have emerged since the birth of digitalisation. It is no longer anywhere near enough to measure citations in the quantitative analysis of publications and there has long been a wealth of serious publications, such as in social media, where the classical indicator canon has ceased to be applicable. This is where alternative metrics, which are explained in detail and their use described in this chapter, come in.

The chapter 'Applications, Practice and Special Issues in Bibliometrics' outlines a colourful bouquet of special examples and prospects of bibliometric research. These include concepts and initiatives, as well as issues concerning the visualisation of and connections between innovation, regional consolidation and trends in bibliometrics.

Bibliometrics would not have been feasible if the data basis of this academic research had not been constantly expanded and optimised. It constitutes a crucial foundation for automatic analysis systems and the processing of vast quantities of data to work in bibliometrics. The chapter 'The Data Basis in Bibliometrics' discusses various commercial and non-commercial databases like those used in bibliometrics, including their pros and cons.

Bibliometrics grew out of the suspicion of being a "sideline" of sociologists, philosophers, mathematicians or librarians. Whilst it has become a core research and application field at various institutions, it has not yet really established itself as a discipline in its own right. In the chapter 'Teaching and Training', we summarise the current state of affairs, where the next generation of scholars can learn from bibliometrics and how the topic is rooted in university curricula.

The handbook concludes with the chapter 'The Future of Bibliometrics', which attempts to provide an outlook from various perspectives as to where bibliometrics might develop alongside the latest developments of academic communication.

Finally, an index should enable core topics and keywords to be found more easily.

1 History and Institutionalization of Bibliometrics

1.1 A Historical Overview of Bibliometrics

Farshid Danesh and Ali Mardani-Nejad

Abstract: Bibliometrics is one of the few methods that analyze data on a large scale. The development and growth of bibliometrics as an area of research have been remarkable in recent decades. However, bibliometrics is not a new phenomenon as it originated from statistical bibliography. The primary purpose of this chapter is to present a historical view of bibliometrics as well as the significant events that took place in this field from the late nineteenth century to the late twentieth century (1870 – 1980). The chapter will explore 116 years to illustrate the advent of bibliometrics, its essential definitions, its pioneers, its theories, and the influential works of this field, from 1873 to the end of 1989.

Keywords: bibliometrics, bibliometrics trend, bibliometric rule, bibliometric theories.

Introduction

As a word, bibliometrics appears to have its roots in "Biblio" and "metric." The term "Biblio" derived from the combined Latin and Greek word "biblion," which is similar to the word Bybel (os), meaning book. Paper derived from the word Byblos, a city of Phoenicia, noted for its export trade in the paper. The word "metrics," on the other hand, indicates the science of measurement, and is derived either from the Latin or Greek word "metricus" or "metrikos," each respectively meaning measurement (Sengupta, 1992). The metric word also specifies the science of measurement, derived from the Greek or Latin word metric, meaning measurement. In this chapter, the historical background of bibliometrics from origin to the formation of theories studied, and the foundations, thoughts, rules, and publications during the eighteen and nineteenth centuries in bibliometrics, is introduced. This historical overview dates back to the early 1870s when the initial idea of bibliometrics was sparked into life in 1873 and progresses to the late 1980s when it matured. This chapter will focus on the first significant bibliometrics events from the 1870s, which lasted more than a century until the end of the 1980s.

Farshid Danesh, Regional Information Center for Science & Technology (RICeST), Shiraz, Iran, farshiddanesh@ricest.ac.ir
Ali Mardani-Nejad, Young Researchers and Elite Club, Najafabad Branch, Islamic Azad University, Najafabad, Iran

https://doi.org/10.1515/9783110646610-003

Fig. 1

The 1870s

1873: The first bibliometrics documentary, entitled "Histoire des Sciences et des Savants Depuis Deux Siècles," was published by French-Swiss botanist Alphonse de Candolle in 1873. He described the changes made in the scientific ability of nations in terms of their membership in scientific communities. His goal in this study was to identify the factors influencing the scientific success of a nation (van Raan, 2004).

The 1890s

1896: Conceptual bibliometrics derived from statistical bibliography and its application dates back to the 1890s. The work of Campbell (1896), which uses statistical methods to study the subject scattering in publications, may be considered as the first attempt in bibliometrics studies (Osareh, 1996).
1896: Pareto, the Italian economist (University of Lausanne), recognized the 80/20 principle in 1896. In a paper entitled "Cours d'économie politique," he published on this principle. Pareto believed that approximately 80 % of the land in Italy was owned by 20 % of the population (Moore, 1897). He then carried out surveys on a variety of other countries and found to his surprise that a similar distribution applied (Backhaus, 1980).

The 1910s

1913: German physicist Felix Auerbach presented German cities' population ranking according to the distribution law. The law that we now refer to as "Zipf's Law" was founded in those years (Auerbach, 1913).

Alfred J. Lotka 1880-1949

Edward Condon 1902-1974

George Kingsley Zipf 1902-1950

Samuel C. Bradford 1878-1948

Fig. 2

1916: French stenographer J.B. Estoup noted that rank (r) and frequency (F) in a French text were related by a "hyperbolic" law, which states that r · F is approximately constant, cf. (Estoup as cited in Harremoës and Topsoe, 2005).
1917: Cole and Eales used a statistical bibliography in 1917 to study citations of comparative anatomical texts from 1550 to 1860 (Cole and Eales, 1917).

The 1920s

1923: Hulme performed a statistical analysis of science history in 1923. His analysis was based on journals that had reached the International Bookkeeping Directory in 17 disciplines. Hulme was the first to invent the term "statistical bibliography" (Hulme, 1923).
1926: Lotka provided a basis for the scientific productivity of the authors. Lotka believed that few authors produce a high percentage of scientific works in scientific subjects. Of course, this does not necessarily reflect the impact and content of these authors' publications (Lotka, 1926; Garfield, 1995).
1927: Gross and Gross also used the citation analysis method for the first time. They used previous periodicals as sources of chemistry citations. Their citation analysis method as a model and sample was used up to 50 years later (Gross and Gross, 1927).
1928: E. Condon, a physicist from the Bell Telephone Company, found regularity in the research on raising the capacity of telephone lines for communication. Condon found that the distribution relationship between lgr & lgnr is close to a straight line AB, with an angle α between line AB and the χ-coordinate. If tgα=γ, then log (rγ • nr) = log K (Qiu et al., 2017).
1929: Zipf defended his Ph.D. dissertation entitled "Relative Frequency as a determinant of phonetic change" (Zipf, 1929).

Fig. 3

The 1930s

1932: Zipf published a book entitled *Selected Studies of the Principle of Relative Frequency in Language* (Zipf, 1932).
1934: Bradford published his articles on scientific journals. He reported that a relatively small number of journals published a high percentage of all papers. Bradford divided journals into two groups: (1) a central core of the journals with their most relevant articles; and (2) several groups around the core (Bradford, 1948).
1934: Paul Otlet, the Belgian librarian, was one of a few people called the Father of Information Science. He added "DOCUMENTATION" to the field of information science, while the Universal Decimal Classification (UDC) was another of his achievements. He also used the term bibliometrics for the first time in the Traité de Documentation book (Rousseau, 2014).
1935: In a book entitled *The Psychobiology of Language: An Introduction to Dynamic Philology,* Zipf stated that the word length is inverse with its relative frequency. This exploration led to the discovery of a law known as Zipf's law, which, in his opinion, and in general, tends to be based on the size of the words, which has an inverse (not necessarily proportional) relation to the number of occurrences (Zipf, 1935).

The 1940s

1948: Herman Fussler characterized the literature of chemistry and physics, the use of "key journals" (Fussler, 1949).
1948: Samuel C. Bradford published "Documentation" in 1948 (Bradford, 1948).
1948: In 1948, Claude Elwood Shannon presented an article entitled "A Mathematical Theory of Communication," which had a striking effect on information science. In this paper, he considers the fundamental problem of communication in reproducing the message sent by the sender to the receiver (Shannon, 1948).
1948: The term "librametrics" was first introduced by Siyali Ramamrita Ranganathan (S.R.R.) in 1948 at the annual Slip Conference in Spain (Ranganathan, 1948).

Fig. 4

1949: Zipf made the first attempt at vocabulary within the text. By studying the abundance of words used in English texts, Zipf found some examples of the principle of least effort (Zipf, 1949).

The 1950s

1955: Eugene Garfield devised a bibliographic system for scientific texts called the "Science Citation Index." The citation index scheme compiled information that was more convenient than the conventional index called "subject index" and could bridge the gap between authors and researchers (Garfield, 1955).
1956: Fano and Kessler introduced the first idea of the bibliographic coupling. The "Bibliographic couple" is one of the bibliometric methods that examine the critical works of authors, documents, and prestigious journals and identifies the relationship between the author's major works and references and essential documents and journals (Sen and Gan, 1983).
1958: Miller and Newman clarified the statistical explanation between the rank and frequency of words in English texts (Miller and Newman, 1958).

The 1960s

1963: Garfield published the Science Citation Index (SCI). The use of this resource has flourished as a tool for various studies and analyses, and today it is one of the most reputable sources for bibliometric studies.
1963: Price published his book entitled *Little Science, Big Science.* For the first time, he used statistical data, "the phenomenon of progressive text growth." He showed that between 1660 and 1960, the number of scientific articles had doubled every 15 years (Price, 1963).
1963: Eugene Garfield and Irving Sher introduced Journals Impact Factor (IF) in the Institute for Scientific Information (ISI) to select journals for the Science Citation Index (Garfield and Sher, 1963; Garfield, 1999). At that time, nobody thought that

the index would be so affected. Now, Impact Factor is an extensive guide to selecting the best journals for the Science Citation Index.
1964: Goffman and Newill presented the "GENERALIZATION OF EPIDEMIC THEORY." They believed that the movement of an idea over time was the same as an infectious virus. The virus grew at a particular time, with some infected, transmitting this idea (TRANSMISSION OF IDEAS), with a feverish peak of awareness in it (Goffman and Newill, 1964). This process may gradually be reduced and improved, leading to the death of thought or ending with its incubation. This model can be used with ISI analyses to predict the prevalence of a research subject, the duration of an outbreak, the number of people affected, and whether an information retrieval system should facilitate the communication of relevant scientific information (Garfield, 1980).
1965: Derek de Solla Price called for the link between articles as the network of scientific articles (Price, 1965).
1967: Leimkuhler presented a mathematical model for explaining the efficiency of publications (Leimkuhler, 1967).
1968: Co-citation of authors was introduced as co-mention by Rosengren in 1968 (Rosengren, 1968). Co-citation studies are the correct representation of the logical mind-body structures of science, and a tool for observing the current state of science and for predicting its future direction.
1968: Merton unveiled a phenomenon called the "Matthew Effect." He believed that the amount of citation received by researchers with an organizational affiliation was as large as the organization, meaning that researchers from large organizations received more citation than those from small organizations (Merton, 1968).
1969: The term Nakometria, later renamed Scientometrics, was first used in the Soviet Union by Vasily Nalimov (Nalimov and Mul'chenko, 1969).
1969: The term "bibliometric" was used in 1969 by Pritchard. He believed that the term bibliography was statistically ambiguous because it may be interpreted as a statistical analysis of bibliographies in statistics. Therefore, the term "bibliometrics" is more effective and comprehensive than the "statistical bibliographic" term and suggests the use of mathematics and statistical methods in the quantitative analysis and aspects of librarianship and information science (Pritchard, 1969).
1969: The term "bibliometrics," as far as is verifiable, was first used in the 1969 issue of *Documentation Journal*. In that issue, the paper by Robert Fairthorne was published, entitled "Bibliometric Description." In its first paragraph, Fairthorne noted that Alan Pritchard was the inventor of the term bibliometrics (Fairthorne, 1969).

The 1970s

1970: By 1970, bibliometrics had become a heading in both Library Literature and Library and Information Science Abstracts (Peritz, 1984, quoted in Hood and Wilson, 2001).

1972: Ortega hypothesis: in an original Spanish language book published in English in 1932, Jose Ortega and Gasset presented a hypothesis that was used, 40 years later, by two American sociologists named Jonathan R. Cole and Stephen Cole, in 1972, in an article entitled "The Ortega Hypothesis," published in *Science* Journal. They examined his hypothesis about physics scientists and concluded that Ortega's hypothesis did not approve in their research populations. However, they did not reject the hypothesis and suggested that it could be considered in other sciences (Cole and Cole, 1972).
1973: The Social Science Citation Index (SSCI) was published at the Institute for Scientific Information (ISI) in Philadelphia, the USA, for social science research areas, under the supervision of Eugene Garfield.
1973: Using methods of group analysis, Small studied co-citation couples. The identification and mapping of the structure of a scientific discipline also became possible through the study of co-citation coupling (Small, 1973).
1975: Moravcsik and Murugesan proposed the first approach to categorizing citations in 1975. Their most important work at that time was weighting citations based on function and context, which had not been carried out until then. They began their studies from 1975 and their first article entitled "Some Results on the Function and Quality of Citations" was published in 1975. They classified scientific papers' citations based on being (a) conceptual or operational, (b) organic or perfunctory, (c) evolutionary or juxtapositional, or (d) confirmatory or negational (Moravcsik and Murugesan, 1975; Jörg, 2008)
1975: Brookes organized the first International Research Forum in Information Science (IRFIS) at University College London to discuss the theoretical aspects of information science (Brookes, 1976).
1976: Price, in his article about a "Cumulative Advantage Distribution," proposed a model in which, statistically, success breeds success in a specific situation. It is common in bibliometric matters and many diverse social phenomena that success seems to breed success. A paper that has been cited many times is more likely to be cited again. In this theory, it would appear that the course of future citation successes is determined statistically by the history of the cited paper, suggesting that citations are generated by a pull mechanism from the previous citation rather than from a push mechanism of the papers that do the citing (Price, 1976).
1976: After publishing the Science Citation Index at the Institute for Scientific Information (ISI), Garfield published citations in order to rank journals and present a regular and analytical report on their status and ranking in a variety of thematic areas and disciplines.
1978: Scientometrics Journal was founded, and the concept of Scientometrics was introduced at the same time in Budapest. According to Bensman, Scientometrics Journal acted as a bridge between the East and the West (Garfield, 2007).
1978: After the publication of two citation indexes of science and social sciences, the Institute for Scientific Information (ISI) found it necessary to view studies in the art

and human sciences using the indexing and organizing method. This led to the arts and humanities citation index being published for the first time (Komatsu, 1999).
1979: Prof. Otto Nacke introduced the term infometric. He defined this concept as the use of mathematical methods to understand phenomena related to information through the description and analysis of these phenomena and the discovery of the rules governing them (Sengupta, 1992).

The 1980s

1981: Based on the analysis of co-citation coupling, White and Griffith presented a new tool for exploring the intellectual structure of science (White and Griffith, 1982).
1983: The co-word analysis, introduced by Callon in 1983, was based on the assumption that the presence of words or concepts together represents the content of that document. By measuring the amount of this co-occurrence, the network can draw the concepts of a scientific field (Callon, 1983).
1983: The Price Award was the first and most prestigious international informetric award, given to individuals with proven and extraordinary achievements in quantitative science studies and applications. Tibor Braun, the founder and editor-in-chief of the Scientometrics journal, created the Price Award in 1983, to keep the memory of Price alive (Derek John De Solla Price Award of the Journal Scientometrics, 2016). One of the goals of this award was to acknowledge the individuals who played a vital role in the development of science.
1984: The first Price medal was awarded to Eugene Garfield in 1984.
1984: In 1984, FID (International Federation for Information and Documentation) set up the Committee on Infometrics to provide valid data for research and development, policymaking, planning, management of institutions, and projects, programs, and scientific activities.
1987: The first international conference on Bibliometrics, Scientometrics, and Infometrics was held in Belgium in 1987 and the second conference in London in 1989.

Conclusion

In this chapter, a historical approach to bibliometrics and a century of essential events were reviewed. From the first efforts of researchers in this field, such as Candolle's attempt to publish the first bibliometrics documentary in 1873, to the studies of White and Griffith and Callon at the beginning of the 1980s for both co-citation and co-word studies, all such scholars published works in this area whose influence would be long-lasting.

The first attempts to record the concept of bibliometrics to the internationalization of this area of human knowledge and the establishment of the world's first scientific organization, named *Institute for Scientific Information (ISI)*, was undertaken by Eugene Garfield (September 16, 1925–February 26, 2017), in particular for citation indexing, abstracting, and citation analysis of reputable papers and publications discussed globally. He also published the first science, social sciences, and arts and humanities citation indexes, while for the ranking of scholarly journals based on valid indicators such as the Impact Factor (IF) he published the first *Journal Citation Report (JCR)*.

Also of note is the publishing of the first specialized journal, *Scientometrics*, the organizing of national and international conferences on the topic of metrics, and quantitative researches and the attention of international communities and associations such as FID and the Committees of Infometrics establishment in FID. All of this resulted in a heightened understanding of the world of science, such as the global importance and necessity of this field as well as its capabilities as an interdisciplinary research area in solving the problems and challenges of research-related areas and macro policymaking, and observing the scientific status of countries, organizations, and scientists at the national and international levels.

References

Auerbach, F 1913, 'Das Gesetz der Bevölkerungskonzentration', *Petermann's Geographische. Mitteilungen*, vol. 59, pp. 7–74.

Backhaus, J 1980, 'The Pareto Principle', *Analyse & Kritik*, vol. 2, no. 2, pp. 146–171.

Bradford, SC 1948, *Documentation*, Crosby Lockwood, London.

Brookes, BC 1976, 'A personal note', in *International Research Forum in Information Science: the Theoretical Basis of Information Science*, pp. 6–7, BLR & D Report 5262, British Library, London.

Callon, M, Courtial, JP, Turner, WA & Bauin, S 1983, 'From translations to problematic networks: an introduction to co-word analysis', *Social Science Information*, vol. 22, no. 2, pp. 191–235.

Campbell, F 1896, *Theory of the National and International Bibliography*, Library Bureau, London.

Cole, FJ & Eales, NB 1917, 'The history of comparative anatomy. Part I: A statistical analysis of the literature', *Sei. Prog*, vol. 11, pp. 578–596.

Cole, JR, & Cole, S 1972, 'The Ortega Hypothesis', *Science*, vol. 178, no. 4059, pp. 368–375.

Derek John de Solla Price Award of the Journal Scientometrics [Website], http://issi-society.org/awards/derek-de-solla-price-memorial-medal/ (July 15, 2020).

Fairthorne, RA 1969, 'Empirical hyperbolic distribution (Bradford-Zipf-Mandelbrot) for bibliometric description', *Journal of Documentation*, vol. 25, no. 4, pp. 319–343.

Fussler, HH 1949, 'The Library Quarterly: Information', *Community. Policy*, vol. 19, no. 1, pp. 19–35.

Garfield, E 1955, 'Citation indexes for science: a new dimension in documentation through association of ideas', *Science*, vol. 122, pp. 108–111.

Garfield, E 1980, 'The epidemiology of knowledge and the spread of scientific information', *Essays of an Information Scientist*, vol. 4, pp. 586–591.

Garfield, E & Sher, IH 1963, 'New factors in the evaluation of scientific literature through citation indexing', *American Documentation*, vol. 14, no. 3, pp. 195–201.

Garfield, E 1995, 'New International Professional Society Signals The Maturing Of Scientometrics And Informetrics', *The Scientist*, vol. 9, no. 16, p. 11.

Garfield, E 1999, 'Journal impact factor: a brief review', *Canadian medical association journal*, vol. 161, no. 8, pp. 979–980.
Garfield, E 2007, 'From the science of science to scientometrics: Visualizing the history of science with histcite software', *Proceedings of ISSI. 1, P. 21–26, 11th International conference of the international society for scientometrics & informetrics, CSIC*, June 25–27, Madrid, Spain.
Goffman, W & Newill, VA 1964, 'Generalization of Epidemic Theory', *Nature*, vol. 204, no. 4955, pp. 255–228.
Gross, PLK & Gross, EM 1927, 'College libraries and chemical education', *Science*, vol. 66, pp. 1229–1234.
Harremoës, P & Topsoe, F 2005, 'Zipf's law, hyperbolic distributions, and entropy loss', *Electronic Notes in Discrete Mathematics*, vol. 21, pp. 315–318.
Hood, WW & Wilson, CS 2001, 'The literature of bibliometrics, scientometrics, and informetrics', *Scientometrics*, vol. 52, no. 2, pp. 291–314.
Hulme, FW 1923, *Statistical Bibliography in relation to the growth of modern civilization*, Grafton, London.
Jörg, B 2008, 'Towards the Nature of Citations', in *Poster Proceedings of the 5th International Conference on Formal Ontology in Information Systems (FOIS 2008)*, http://www.dfki.de/~brigitte/publications/FOIS08_Poster_BrigitteJoerg.pdf (July 15, 2020).
Komatsu, S 1999, 'Arts & Humanities Citation Index', *Journal of Information Processing and Management*, vol. 41, no. 12, pp. 989–997.
Leimkuhler, FF 1967, 'The Bradford distribution', *Journal of documentation*, vol. 23, no. 3, pp.197–207.
Lotka, AJ 1926, 'The frequency distribution of scientific productivity', *Journal of the Washington Academy of Science*, vol. 16, no. 12, pp. 317–323.
Merton, RK 1968, 'The Matthew effect in science', *Science*, vol. 159, no. 3810, pp. 56–63.
Miller, GA & Newman, EB 1958, 'Tests of a statistical explanation of the rank-frequency relation for words in written English', *American Journal of Psychology*, vol. 71, pp. 209–218.
Moore, HL 1897, 'Cours d'Économie Politique', by V. Pareto, Professeur à l' Université de Lausanne, vol. I, p. 430, I896; vol. II. p. 426, I897; Lausanne: F. Rouge, *The Annals of the American Academy of Political and Social Science*, vol. 9, no. 3, pp. 128–131.
Moravcsik, MJ & Murugesan, P 1975, 'Some Results on the Function and Quality of Citations', *Social Studies of Science*, vol. 5, no. 1, pp. 86–92.
Nalimov, VV & Mul'chenko, ZM 1969, *Naukometriya. Izuchenie nauki kak informatsionnogo protsessa (scientometrics: Study of science as an information process.)*, p. 192, Nauka, Moscow.
Ortega, JY y Gasset 1932, *The Revolt of the Masses*, pp. 84–85, Norton, New York.
Osareh, F 1996, 'Bibliometrics, Citation Analysis, and Co-Citation Analysis: A Review of Literature I', *Libri*, vol. 46, pp. 149–158.
Price, DJdS 1965, 'Networks of Scientific Papers', *Science*, vol. 149, no. 3683, p. 515.
Price, DJdS 1963, *Little Science, Big Science*, Columbia University Press, New York.
Pritchard, A 1969, 'Statistical Bibliography or Bibliometrics', *Journal of Documentation*, vol. 25, pp. 348–349.
Price, DJdS 1976, 'A General Theory of Bibliometric and other Cumulative Advantage Processes', *Journal of the American Society for Information Science*, vol. 27, no. 5, pp. 292–306.
Qiu, J, Zhao, R, Yang, S & Dong, K 2017, *Informetrics: Theory, methods, and applications*, pp. 121–143, Springer, Singapore.
Ranganathan, SR 1948, Proc. of the Aslib's Ann. Conf., Leamington Spa, Great Britain.
Rosengren, KE 1968, *Sociological aspects of the literary system*, Nature och Kultur, Stockholm.
Rousseau, R 2014, 'Forgotten founder of bibliometrics', *Nature*, vol. 510, p. 218.

Sen, SS & Gan, SS 1983, ‘A Mathematical Extension Of The idea of bibliographic coupling and its application’, *Annals of library science and documentation*, vol. 30, no. 2, pp. 78–82.

Sengupta, IN 1992, ‘Bibliometrics, Informetrics, Scientometrics, and bibliometrics. An Overview’, *Libri*, vol. 42, no. 2, pp. 75–79.

Shannon, CE 1948, ‘A Mathematical Theory of Communication’, *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423.

Small, H 1973, ‘Co-citation in the scientific literature: A new measure of the relationship between two documents’, *Journal of the American Society for information science banner*, vol. 24, no. 4, pp. 256–269.

van Raan, AF 2004, ‘Measuring science: Capita Selecta of current main issues’, in H. Moed, W. Glänzel & U. Schmoch, *Handbook of quantitative science and technology research: the use of publication and patent statistics in studies of S&T systems*, Kluwer, Dordrecht.

White, HD & Griffith, BC 1982, ‘Authors as markers of intellectual space: Co-citation in studies of science, technology, and society’, *Journal of documentation*, vol. 38, no. 4, pp. 227–255.

Zipf, G. 1949, *Human behavior and the principle of least effort*, Addison Wesley, Boston.

Zipf, GK 1929, ‘Relative frequency as a determinant of phonetic change’, *Harvard Studies in Classical Philology*, vol. 40, pp. 1–95.

Zipf, GK 1932, *Selected studies of the principle of relative frequency in language*, Harvard University Press, Cambridge, Mass.

Zipf, GK 1935, *The psycho-biology of language*, Houghton Mifflin, Oxford, England.

1.2 Institutionalization and Professionalization of Bibliometrics

Niels Taubert

Abstract: This article offers an overview as to what extent bibliometrics is an institutionalized field of research and to what extent it is a professional practice. After different dimensions of institutionalization (stabilization of a cognitive core, evolvement of a social structure, development of a communication network) have been reviewed an overview of the dimension of professionalization is given. These include the formation of a specific type of expertise and the establishment of a claim for jurisdiction for a social problem. The article concludes with the following diagnosis: although bibliometics has been established as a field of research, the process of professionalization has to be characterized as incomplete. The reason for this has both to do with missing tracks of formal qualifications together with the proliferation of automated bibliometric tools to non-experts.

Keywords: bibliometrics, communication system, professionalization, institutionalization, infrastructures, professional code.

Introduction

This article offers both an overview as to what extent bibliometrics is an institutionalized field of research and to what extent it is a professional practice. It therefore builds upon two sociological concepts: from a viewpoint of the sociology of the professions (Abboth, 1988; Petersohn, 2016; Petersohn and Heinze, 2018), professions share some common features. First, they comprise a cognitive core of a specific expertise that is usually based on academic knowledge. Second, they address a societally relevant problem for which they put forward cognitive and social claims for jurisdiction and control. Third, such claims are often contested by competing groups. Fourth, expertise is institutionalized in the form of professionals, organizations as well as commodities such as databases, classification, and other artifacts.

The institutionalization of the cognitive core of the profession in which academic knowledge is developed can be described with reference to the sociology of science. Within this tradition it is commonplace that the epistemic process of the establishment of a field of knowledge goes hand in hand with the formation of a social structure. An institutionalized field of knowledge therefore finds its expression in research institutes and groups dedicated to the field that forms a community of scholars and are connected by a network of communication. The community is reproduced by a specific (academic) training which introduces future members to the field. Institu-

Dr. **Niels Taubert**, head of the working group Bibliometrics and a sociologist of science, technology, and digital media by training, niels.taubert@uni-bielefeld.de

https://doi.org/10.1515/9783110646610-004

tionalized communities are organized by scholarly societies (Weingart and Schwechheimer, 2007) and share common criteria of excellence for the evaluation of contributions to the knowledge. If a professional field of application exists, it is likely that professional norms or ethics are codified.

In what follows, the complementary dimensions of the two concepts mentioned above are used as a heuristics to describe the degree of institutionalization and professionalization of bibliometrics. In drawing a multifaceted picture, it will be shown that bibliometrics today is an institutionalized field of research in which the reproduction of the social structure of the bibliometrics community in terms of training is lacking. Because of the absence of a formal qualification together with the availability of bibliometric tools for non-experts, the claim for jurisdiction and control is contested, especially in the field of research evaluation. With respect to control, the degree of professionalization tends to be low.

Cognitive Core, Communication System, and Social Structure in Academic Bibliometrics

Bibliometrics, i.e., methods that deal with the analysis of large numbers of scientific publications and the relations between them – the citations –, has a long tradition. Although the term bibliometrics was invented by Alan Prichard in 1969, bibliometric studies are much older. Besides the seminal and well known works of Alfred Lotka (1926) on the distribution of productivity and of Derek de Solla Price (1963) on the growth of science, one should note that systematic studies on the development of publications and citations have been conducted in psychology (Godin, 2006) as well as in chemistry (Gross and Gross, 1927). Since these first bibliometric studies were based on manual collections of data, the availability of a database that is continuously updated and can be reused for different purposes is of critical relevance. Therefore, the Science Citation Index (published first in 1963 for the year 1961) played a major role for the institutionalization of bibliometric research. The consolidation of the cognitive core of the field began during the 1970s with a large expansion since the 2000s (Yang et al., 2016; Petersohn and Heinze, 2018). The field has developed and a diverse range of topics have been studied, including, for example, scientific productivity, citation networks, co-operations, disciplinary profiles of institutions, mobility and career of researchers, performance indicators, and so on. Regarding its location in the scientific landscape, it is often stated that it has been differentiated as a subfield in the library and information sciences (LIS) but it also holds strong links to the social studies of science.

The institutionalization of a cognitive core of bibliometrics went hand in hand with the foundation of communication channels for the dissemination of research in the field. For years there was no alternative to *Scientometrics* (founded in 1978) and the *Journal of the American Society for Information Science and Technology*. This changed in 1991, when *Research Evaluation* was first published. During this period, the publication output of *Scientometrics* grew enormously: the journal started with five issues in one volume, changed to six issues in two volumes in 1986, then published nine issues in three volumes, and ended up with 12 issues in four volumes on an annual basis from 2005 onwards. The increase of the publication output of the bibliometrics community is also reflected by the foundation of new journals such as the *Journal of Informetrics* and *Collnet Journal of Scientometrics and Information Management*, which both appeared from 2007 onwards, and the *Journal of Scientometric Research and Publications*, which started in 2012. With respect to institutionalization, it is interesting to note that there is still no journal that is exclusively dedicated to bibliometrics. The main journals have a wider scope and address a broader community as they refer to more comprehensive terms like "scientometrics" or "quantitative analysis of science" in their self-descriptions (aims and scope). Besides these core-journals of the field, there are a large number of channels with a more general scope that are also open to contributions from bibliometrics. These journals usually cover library and information studies, documentation, management and use of information technologies or librarianship in general.

Like other fields of research, the bibliometrics community has also developed institutionalized forms of face to face exchanges in conference series. The oldest series is the yearly conference of the *International Society for Scientometrics and Informetrics* (ISSI), first conducted in 1987. A second conference series is the Science and Technology Indicators (STI) Conference that was initiated in 1988 by the Lisbon Institute. The community is represented by two major scholarly societies, the *International Society for Scientometrics and Informetrics* and the *American Society for Information Science.*

Bibliometric research is conducted at many places. Among important locations are the Leiden Centre for Science and Technology Studies (CWTS), the Centre for R&D Monitoring at KU Leuven, the Scimago Lab in Spain, the Karolinska Institute in Sweden, and the contract research company Science Metrics in Canada. In Germany, institutions that work in the field of bibliometrics form a network connected by the Competence Centre for Bibliometrics. In some contrast to the cognitive core, the established institutions, and the communication channels, the institutionalization of the reproduction function of the community is rather low. Bibliometric competencies are often developed in workshops run by the database providers at CWTS (Cox et al., 2019) or at the *European Summer School on Scientometrics* (ESSS) in combination with training on the job. There is some evidence that the inclusion of bibliometrics in the curricula of LIS programs is weak (Corrall, Kennan, and Afzal, 2013; Petersohn, 2016; Cox et al., 2019).

The institutionalization of academic bibliometrics can be summarized as follows: there is an epistemic core that has been established on the scientific landscape with paradigmatic "example" works and a clear-cut methodology together with institutes that – among others – conduct research in bibliometrics. Institutionalized channels for the communication of research also exist, as do scholarly societies that organize the community. The only shortfall in the institutionalization is that there is a considerable lack of formal ways of qualifications and a lack of certificates signaling qualification in bibliometrics to clients interested in expertise.

Infrastructures for Bibliometrics

Citation databases act as important infrastructures, both for bibliometric research as well as for professional practice. Therefore, their development is sketched in the next step. For more than 40 years the databases of the Institute for Science Information – at the beginning the Science Citation Index only, later also including the Social Science Citation Index, the Arts and Humanities Citation Index, the Conference Proceedings Citation Index, and the Book Citation Index – virtually held a monopoly for the provision of data for citation analysis. From 2000 onwards, a noticeable diversification of sources for citation information has taken place: Elsevier launched Scopus in 2004, Google published Google scholar in 2004 and the Digital Science (part of the Holtzbrinck Publishing Group) came up with Dimensions in 2018. Aside from the large interdisciplinary databases, subject-specific databases such as, for example, the Astrophysics Data System (ADS) (Eichhorn, 2004), zbMATH (Teschke, 2015), and MathSciNet (Davis and Fromerth, 2007) also include citation information as well as academic social network sites like ResearchGate and Academics (both launched in 2008). The increase in the number of sources of citation information is based on digital publishing and the possibility to automatically collect, process, and extract publication and citation information (de Rijcke and Rushforth, 2015), allowing a new regime to measure science that can be called “Quantification 2.0” (Krüger, 2020).

In addition, proprietary products like SciVal and InCite as well as free bibliometric tools like Publish or Perish (Harzing, 2010) and Google Scholar Metrics (Jascó, 2012) were released. What is momentous about this development is that they provide easy-to-use instant bibliometric analysis and off-the-shelf indicators that can be used by anyone, including non-bibliometric experts like research administrators or scientists themselves.

Bibliometrics and its References to Societally Relevant Problems

From its very beginning the problems to which bibliometrics should respond to were manifold. First, there was an interest in understanding the dynamics of science as described above. Second, there was (and still is) a perceived flood of scientific publications resulting in an information crisis (Wouters, 2000, p.66). As far as bibliometrics is concerned, the problem has at least two dimensions. On the one hand, there is the problem of decision-making for librarians, about which journals should be included in library collections, especially if funds and physical space are limited. On the other hand, there is the problem for scientists to find and select interesting and relevant publications, especially if time for reading is limited. Following the initial conception of the Science Citation Index, this was the problem Eugene Garfield wanted to create a solution for. The Science Citation Index was understood as an association-of-ideas-index (Garfield, 1955) and should help scientists to track the take-up of scientific ideas and research in other publications. The Journal Citation Report (JCR) introduced in 1975 refers to the second dimension – the decision-making problem of libraries – and provides information for selecting journals.

From the 1980s onwards, a third problem originated that is today most strongly associated with bibliometrics: to inform external stakeholders about the performance of different entities of the research system and to inform decision-makers about research funding. In this context of application, bibliometrics is used to measure scientific "productivity" and "impact". Evaluative bibliometrics is applied on different levels and at different occasions. Besides formal research evaluation in the context of performance-based funding systems (Butler, 2010) on the national level, evaluative bibliometrics is also used within research organizations for monitoring or informed decision-making and also on the level of individual researchers (e. g., appointment committees, tenure decisions, application for third party funding). Today, the allocation problem can be regarded as the most important societal problem that bibliometrics refers to (Taubert, 2013). Nevertheless, the jurisdictional claim for professional control in that field is by no means uncontested (Petersohn, 2016; Petersohn and Heinze, 2018). First, it competes with the mechanisms of self-regulation of science that the relevance and impact can only be assessed by peer scientists of the field. Second, it competes with non-trained groups using tools that offer instant bibliometric analysis. In many cases this kind of use does not thereby meet the professional standards, and has been criticized as not being well informed and being ill applied (Gingras, 2014). With respect to the extent of control over the field of research evaluation, the degree of professionalization of evaluative bibliometrics tends to be low.

Professional Code as a Reaction to Incomplete Professionalization

The contestation of bibliometrics evaluation can also be observed in different resolutions that refer to this type of use. While the statement of the board of directors of the IEEE (2014) refers to the malpractice in the use of bibliometric indicators (and supports calls for professionalization), The San Francisco Declaration on Research Assessment (DORA, 2012) addresses both (possible) extensive claims for jurisdiction of evaluative bibliometrics as well as malpractice. It likewise calls for not using the Journal Impact Factor as a surrogate for the measurement of individual articles (malpractice) and for assessing the quality of a contribution on the ground of the scientific content and not on metrics (restriction of extensive claims of evaluative bibliometrics).

The bibliometrics community also reacted to the contested nature of their claim with a resolution that can be regarded as the codification of professional norms. In its ten principles, it self-restricts the role of bibliometric analysis in the context of research evaluation, on the one hand, by stating that bibliometric analysis should support but not substitute expert assessment. The control over the field of evaluation should be exercised by both the self-regulating mechanisms of science (qualitative assessment by peers) and evaluate bibliometrics. On the other hand, the claim for jurisdiction is defended against groups without competencies in bibliometrics by the formulation of standards (e. g. careful selection of indicators that measure the performance against the mission of the entity that is evaluated, selection of field adequate indicators taking different publication and citations practices into account, regular updating of indicators (Hicks et al. 2015)). Such demands exclude research evaluation based on automated tools for bibliometric analysis.

Missing tracks of formal qualifications in bibliometrics and missing certificates that could signal competencies together with the proliferation of automated bibliometric tools to non-experts make it likely that claims of jurisdiction on research evaluation will also be contested in the future.

References

Abboth, A 1988, *The System of Professions. An Essay on the Division of Expert Labor*, The University of Chicago Press, Chicago and London.

Butler, L 2010, 'Overview of models of performance based research funding systems', in OECD (ed.), Performance-based Funding for Public Research in Tertiary Education Institutions, Workshop Proceedings, pp. 23 – 52, OECD Publishing.

Corrall, S, Kennan, MA & Afzal, W 2013, 'Bibliometrics and Research Data Management Services: Emerging Trends in Library Support for Research', *Library Trends*, vol. 61, no. 3, pp. 636 – 674.

Cox, A, Gadd, E, Petersohn, S & Sbaffi, L 2019, 'Competencies for bibliometrics', *Journal of Libarianship and Information Science*, vol. 51, no. 3, pp. 746 – 762.

Davis, PM & Fromerth, MJ 2007, 'Does the arXiv lead to higher citations and reduced publisher downloads for mathematics articles?', *Scientometrics*, vol. 71, no. 2, pp. 203 – 215.

De Solla Price, D 1963, *Little Science, Big Science*, Columbia University Press, New York et al.

De Rijcke, S & Rushforth, A 2015, 'To intervene or not to intervene; is that the question? On the role of scientometrics in research evaluation', *Journal of the Association for Information Science and Technology*, vol. 66, no. 9, pp. 1954 – 1958.

DORA 2012, *San Francisco Declaration on Research Assessment*, https://sfdora.org/read/ (July 15, 2020) Fehler! Hyperlink-Referenz ungültig..

Eichhorn, G 2004, 'Ten Years of the Astrophysics Data System', *Astronomy and Geophysics*, vol. 45, pp. 3.7 – 3.9.

Garfield, E 1955, 'Citation Indexes for Science. A new dimension in documentation through association of Ideas', *Science*, vol. 122, no. 3159, pp. 108 – 111.

Gingras, Y 2014, *Bibliometrics and Research Evaluation. Uses and Abuses*, The MIT Press, Cambridge (Mass.) et al.

Godin, B 2006, 'On the origins of bibliometrics', *Scientometrics*, vol. 68, no. 1, pp. 119 – 133.

Gross, PLK & Gross, EM 1927, 'College Libraries and Chemical Education', *Science*, vol. 66, no. 1713, pp. 385 – 389.
Harzing, AW 2010, *Publish or Perish. Your guide to effective and responsible citation analysis*, Tarma Software Research, Melbourne.
Hicks, D, Wouters, P, Waltman, L, de Rijcke, S & Rafols, I 2015, 'Bibliometrics: The Leiden Manifesto for research metrics', *Nature*, vol. 520, no. 7548, pp. 429 – 431.
IEEE 2014, 'Appropriate Use of Bibliometric Indicators for the Assessment of Journals, Research Proposals, and Individuals', *IEEE Computer Graphics and Application*, vol. 34, no. 2, pp. 86 – 87.
Jascó, P 2012, 'Google Scholar Metrics for Publications. The software and content features of a new open access bibliometric service', *Online Information Review*, vol. 36, no. 4, pp. 604 – 619.
Krüger, AK 2020, 'Quantification 2.0? Biblilometric Infrastructures in Academic Evaluation', *Politics and Governance*, vol. 8, no. 2, forthcoming.
Lotka, AJ 1926, 'The frequency distribution of scientific productivity', *Journal of the Washington Academy of Sciences*, vol. 16, no. 12, pp. 317 – 323.
Petersohn, S 2016, 'Professional competencies and jurisdictional claims in evaluative bibliometrics: The educations mandate of academic librarians', *Education for Information*, vol. 32, no. 2, pp. 165 – 193.
Petersohn, S & Heinze, T 2018, 'Professionalization of bibliometric research assessment. Insights from the history of the Leiden Centre for Science and Technology Studies (CWTS)', *Science and Public Policy*, vol. 45, no. 4, pp. 565 – 578.
Prichard, A 1969, 'Statistical Bibliography or Bibliometrics?' *Journal of Documentation*, vol. 25, no. 4, pp 348 – 349.
Taubert, N 2013, 'Bibliometrie in der Forschungsevaluation. Zur Konstitution und Funktionslogik wechselseitiger Beobachtung zwischen Wissenschaft und Politik', in JH Passoth & J Wehner (eds.), Quoten, Kurven und Profile. Zur Vermessung der sozialen Welt, pp. 197 – 204, Springer VS, Wiesbaden.
Weingart, P & Schwechheimer, H 2007, 'Institutionelle Verschiebungen der Wissensproduktion – Zum Wandel der Struktur wissenschaftlicher Disziplinen', in Weingart, P (ed.), Nachrichten aus der Wissensgesellschaft. Analysen zur Veränderung der Wissenschaft, pp. 41 – 54, Velbrück Wissenschaft.
Wouters, P 2000, 'Garfield as alchemist', in B Cronin & HB Atkins (eds.), The Web of Knowledge, pp. 65 – 71, Information today, Medford.
Yang, S, Han, R, Wolfram, D & Zhao, Y 2016, 'Visualizing the intellectual structure of information science (2006 – 2015): Introducing author keyword coupling analysis', *Journal of Informetrics*, vol. 10, no. 1, pp. 132 – 150.

1.3 Eugene Garfield and the Institute for Scientific Information

David A. Pendlebury

Abstract: Eugene Garfield and his Institution for Scientific Information (ISI) helped revolutionize information retrieval in the second half of the twentieth century by introducing the concept of citation indexing for scientific literature. *Science Citation Index* data also served as a foundation for quantitative studies in the history and sociology of science and eventually gave birth to the field of scientometrics. From its founding in 1960 until its sale to Thomson Corporation in 1992, ISI introduced a range of current awareness and information retrieval products and services covering the literature of the sciences, social sciences, and humanities. Other products, such as the *Journal Citation Reports*, permitted analysis of citations as measures of communication and research performance. ISI colleague Henry Small, with the support of Garfield, introduced science mapping in the 1970s to reveal the socio-cognitive structure of research. Clarivate Analytics acquired the ISI product range from Thomson Reuters Corporation in 2016 and today continues the original business and intellectual legacy of Garfield.

Keywords: citation analysis, citation indexing, Clarivate Analytics, *Current Contents*, Eugene Garfield, history and sociology of science, impact factor, Institute for Scientific Information (ISI), *Science Citation Index* (*SCI*), *Web of Science* (*WoS*).

Eugene Garfield (1925 – 2017) established the Institute for Scientific Information (ISI) in Philadelphia, Pennsylvania, USA, in 1960 as a commercial entity to produce a wide range of current awareness and information retrieval products, including *Current Contents* and the *Science Citation Index*. ISI (1960 – 1992) proved to be an innovative and risk-taking organization that pioneered many new concepts while rapidly adopting the latest technology in information processing, storage, and methods of information dissemination (Lazerow, 1974; Cawkell and Garfield, 2001; Lawlor, 2014). Its products were progressively transitioned from print to magnetic tape, to diskette, to CD-ROM, and to the World Wide Web. For more than three decades, the man and the business were inseparable in the eyes of customers and the research community at large.

David A. Pendlebury, Institute for Scientific Information, Clarivate Analytics, 160 Blackfriars Road, London SE1 8EZ, United Kingdom, david.pendlebury@clarivate.com

https://doi.org/10.1515/9783110646610-005

The Path to an Idea

The notion of a citation index for the sciences upended traditional ideas of indexing and information retrieval, which were previously focused on subject indexing by professionally trained personnel using controlled vocabularies. In the 1950s the increase in the production of scientific information put great strain on traditional indexing methods, resulting in delays of several years from publication to availability of the information in an index.

At the beginning of the decade, in 1951, after obtaining a B.S. in Chemistry at Columbia University in 1949 and working in the lab of physical chemist Louis P. Hammett, Garfield joined the Welch Medical Indexing Project at Johns Hopkins University, funded by the Army Medical Library, the predecessor of the National Library of Medicine. He explored machine methods to speed indexing of the scientific literature. In 1953 he organized the first symposium on this subject. It was after this that Garfield learned about citation indexing employed in the legal profession in *Shepard's Citations*. This index, in existence since 1873, served lawyers and judges by showing which legal decisions had been subsequently affirmed, overturned, or modified. Although Garfield made significant contributions to advancing traditional indexing using the then new IBM 101 Electronic Statistical Machine, his discovery of citation indexing was transformative in his thinking about automating indexing and improving information retrieval.

He immediately grasped the benefits of applying this method to the scientific literature: by recording references given by authors in their papers, one could reliably and precisely find related publications and, moreover, search "forward" to identify later papers that cited an earlier one of interest. After receiving his M.S. degree in library science from Columbia University in 1954, Garfield published a paper in *Science* in 1955 proposing a citation index for the sciences, which he called an "association-of-ideas" index and the researchers supplying the citations his "army of indexers." Citation indexing did not require professional indexing using subject headings and controlled vocabulary and was also independent of changing terminology. Moreover, citation indexing, he stated, broke the "subject index barrier" of conventional monodisciplinary indexes since citations could and did point to literature outside the boundaries of any defined coverage (Garfield, 1955; Weinstock, 1971; Garfield, 1979; McVeigh, 2017; de Araújo, 2019).

The *Science Citation Index*

So unconventional was Garfield's proposal that he met with indifference or opposition. In 1959, Joshua Lederberg, the 1958 Nobel laureate in Physiology or Medicine, wondered what happened to Garfield's proposal of 1955 and contacted him to ask about the development of a citation index, for which he expressed support. Lederberg's advocacy proved essential and a prototype focusing on genetics (*Genetics Ci-*

tation Index), with support from the US National Institutes of Health (NIH), was developed in 1962 and 1963 (Lederberg 1963). Garfield's suggestion to NIH and the National Science Foundation (NSF) to publish the full *Science Citation Index* (the name was Lederberg's suggestion) was not adopted, so Garfield took the risk to produce this himself in 1963 (Wouters, 1999; De Bellis, 2009).

The next year saw the first commercial version of the *Science Citation Index (SCI)*, issued in quarterly fascicles (Garfield, 1964). Garfield traveled the world to spread the gospel of citation indexing. Institutional subscriptions steadily increased. In 1973, he introduced the *Social Sciences Citation Index (SSCI)* and, in 1978, added the *Arts & Humanities Citation Index (A&HCI)* . The publication of the citation indexes in book form made for challenging manual searching as one consulted a Source Index of items listed by first author, a Permuterm Subject Index (from 1966 onwards) which listed pairs of title words, and a Citation Index – each section an entry point to, respectively, indexed papers, subjects and subtopics, and cited works and their citing papers. Online access via dial-up vendor services such as Dialog in the 1970s ameliorated manual searching through multiple and heavy printed volumes. A CD-ROM version of the *SCI* was introduced in 1988. Following the sale of ISI to the Thomson Corporation in 1992, a web version, bringing together the *SCI*, *SSCI*, and *A&HCI* and named the *Web of Science* (*WoS*), appeared in 1997 (Schnell, 2018).

As of the beginning of 2020, and under the ownership of Clarivate Analytics (since 2016), the *WoS* Core Collection indexed more than 21,000 journals and, in total, included about 77 million source items and some 1.5 billion cited references. Coverage for the *SCI* and *SSCI* extends back to 1900, whereas the *A&HCI* spans from 1975 to present. In addition to the flagship databases, *WoS* also includes the *Conference Proceedings Citation Index* (1990-present), introduced in 2008; the *Book Citation Index*, (2005-present), introduced in 2011; and, the *Data Citation Index*, introduced in 2012. The *Emerging Sources Citation Index* (2005-present), which focuses on journals of regional importance and in emerging fields, was added to *WoS* in 2015.

The Journal Impact Factor

In the early 1960s, Garfield and his ISI colleague Irving H. Sher (1924–1996) examined patterns of journal-to-journal citations to identify the most-cited journals in each field (Garfield, 2000a). These data fed back into decisions about the journals to be indexed in the *SCI* if they were not already. Garfield and Sher formulated an indicator of journal importance and influence: the impact factor (Garfield and Sher, 1963; Garfield, 1972; Garfield, 1976; Bensman, 2007; Archambault, 2009; Pendlebury and Adams, 2012; Larivière and Sugimoto, 2019). The impact factor was first published in the *Journal Citation Reports* (*JCR*), a volume appended to the 1975 *SCI*. The *JCR* has always included a wealth of information apart from the impact factor, including total papers, total citations, self-citations, citing and cited journal data, cited and citing half-life data, an immediacy index, and other statistics. In 1990 the *JCR* became a separate product. It offered two editions, for the sciences and for the social sciences, and over the years has been delivered in print, microfiche, CD-ROM, and on the web.

The journal impact factor became one of the most popular but most controversial of the company's offerings. What was designed to aid ISI itself in journal selection, librarians in collection development, and information scientists and other scholars in understanding scientific and scholarly communication patterns among journals turned into a proxy for the quality of individual papers published in a journal and a surrogate measure of performance in the evaluation of individuals. Publication in a high-impact journal, however, does not mean a paper or person will be highly cited, owing to the characteristic skewed distribution of citation statistics. Garfield pointed out the misuse, in public and print, but was nevertheless frequently blamed for applications of the impact factor that were never intended (Garfield, 2006).

Current Contents

While at Johns Hopkins in the early 1950s, Garfield created *Contents in Advance*, a publication that reproduced tables of contents of journals in library and information science. He produced it for his own interests and needs but offered it to others as a product. This was the origin of *Current Contents*. His solution to the problem of current awareness was "ridiculously simple" (Grimwade, 2018): provide in compact form a listing of just-published journal literature to alert researchers to articles of interest even before journals arrived on library shelves. In 1955 he started a contents-page service for management science which he titled *Management's DocuMation Preview*. It was later renamed *Current Contents in Management* but failed to earn a market. In 1956, and related to consulting work for pharmaceutical companies, Garfield started *Current Contents of Pharmaco-Medical, Chemical & Life Sciences*. *Current Contents of Space, Electronic & Physical Sciences Including Pure & Applied Chemistry* appeared in 1960, and thereafter the *Current Contents* business took off. In fact, it was the success of *Current Contents* that allowed Garfield to underwrite the development of the *SCI* and other products, such as *Index Chemicus*, in the 1960s. Editions, under various names, appeared and merged over the years (Lawlor, 2014; Grimwade, 2018). Eventually, there were seven editions of *Current Contents: Life Sciences*; *Clinical Medicine*; *Physical, Chemical & Earth Sciences*; *Agriculture, Biology & Environmental Sciences*; *Engineering, Computing & Technology*; *Social and Behavioral Sciences*; and *Arts & Humanities*.

One appeal of *Current Contents* was Garfield's own writing in the form of essays appearing sporadically in the 1960s but weekly from 1972. His essays, and other materials, are collected in 15 volumes (Garfield, 1977–1993) and available on his personal website (http://www.garfield.library.upenn.edu/.

Selective Dissemination of Information and Document Delivery

Another area of activity for Garfield and Sher in the 1960s dealt with what was then known as Selective Dissemination of Information (SDI) services. Drawing on the data recorded for *SCI* and *Current Contents*, a new service was introduced in 1965: *Automatic Subject Citation Alert* (*ASCA*). With *ASCA*, a researcher would create a search profile including any combination of author names, title words, organization names, journals, and, importantly, cited references that represented the researcher's interests. The search profile was run against the ISI data weekly and a printed report was then mailed to the researcher (Garfield and Sher, 1967).

In the days before transmission of digital forms of journals, ISI received multiple copies of the journals it indexed. These copies were cannibalized to provide copies of specific articles requested by ISI customers, and a royalty would be paid to the publisher. This document delivery service, launched in the 1960s, was originally called *OATS*, for *Original Article Tears Sheet Service*, but was later called *The Genuine Article*.

Chemical Information Products

Educated in chemistry, Garfield took interest at the Welch Medical Indexing Project in "chemical nomenclature used in Medical Subject Headings (MESH) and understanding the need for new approaches to retrieving chemical information" (Garfield, 2000b). Later, in 1958, working as a consultant to the Pharmaceutical Manufacturers' Association, Garfield indexed steroid compounds in the literature and recognized the possibility of algorithmically locating and identifying newly reported compounds. At this time, he invented a method for converting chemical names to molecular formulas, earning a Ph.D. in structural linguistics from the University of Pennsylvania for this work in 1961. He introduced *Index Chemicus* in 1960 as the first product of ISI. It reported new compounds and included graphical abstracts along with chemical formulas. It eventually became *Current Abstracts of Chemistry and Index Chemicus* in 1970. A range of chemical information products followed *Index Chemicus: Index Chemicus Registry System*, in 1968, which permitted computer searching of chemical structures using Wiswesser Line Notation; *Chemical Substructure Index*, in 1971; *Automatic New Structure Alert*, an SDI service, in 1973; *Current Chemical Reactions*, in 1979; and *Reaction Citation Index*, in 1995 (Garfield, 2001). These products were offered in a variety of media as technology changed through the years. The *WoS* Core Collection today includes *Current Chemical Reactions* (since 1985) and *Index Chemicus* (since 1993).

The History, Sociology, and Structure of Science: Derek Price, Robert Merton, and Henry Small

In his 1955 paper Garfield emphasized the importance of citation indexing for information retrieval but noted the value of citation data for the historian, both to reveal the transmission of ideas and the extent that a paper had been cited (Garfield, 1955). He returned to the theme in 1963 and expressed his interest in computer-generated "topological network diagrams which show the chronological and derivational relationships between scientific papers and therefore scientific discoveries" (Garfield, 1963). This paper references personal communication the previous year with two individuals who would figure largely in the use of ISI's citation data for historical and sociological research: Derek J. de Solla Price (1922–1983) of Yale University and Robert K. Merton (1910–2003) of Columbia University.

Price, a physicist by training and historian of science, had published *Science Since Babylon* and would publish *Little Science, Big Science* the year after his contact with Garfield (Price, 1961; Price, 1963). In both he demonstrated his interest in using the "tools of science on science itself." Once made aware of the new *Science Citation Index*, he asked for and received data from Garfield related to his interests, including the nature of the research front which he described as a growing "epidermal layer" of papers (Price, 1965) and cumulative advantage processes in citation that contribute to characteristic skewed distributions (Price, 1976). Price laid a foundation in quantitative studies of science using ISI citation data and in doing so helped establish the field of scientometrics. He also advocated strongly for the use of publication and citation analysis in planning, policymaking, and funding. When the NSF produced its first *Science Indicators* report issued in 1973, it included *SCI* data on national publication output and citation impact.

Merton, the leading sociologist of science, also lent support to Garfield and ISI by describing citation as part of the normative behavior of scientists with respect to acknowledging intellectual property rights. He called the citation a "pellet of peer recognition" and noted the moral imperative to cite one's colleagues in the repayment of intellectual debts (Merton, 1988). Harriet Zuckerman, a student of Merton, has summarized the influence of Garfield and citation analysis in the sociology in science (Zuckerman, 2018).

Henry Small (1941-), an historian of science, arrived at ISI in 1972. He soon introduced the technique of co-citation clustering to define specialty research areas (Small, 1973). The next year, with Belver Griffith (1931–1999), he demonstrated science mapping of the literature representing its socio-cognitive structure as determined by researchers themselves through their patterns of citation. In the following years, Small extended and improved techniques of science mapping, studied the validity of the maps in relation to expert opinion, and explored the connections from one realm to another, noting how these links are "threads that hold the fabric of science together" (Small and Garfield, 1985). Price welcomed Small's work in mapping science, calling it "revolutionary in its implications" and a step toward defining a natural order of research that could be used to create a "giant atlas of the corpus of scientific papers that can be maintained in real time for classifying and monitoring developments as they occur" (Price, 1980). In the 1980s, ISI introduced the *Atlas of Science*, first in book form and thereafter as a series of review journals (Grimwade, 2018). In retrospect, these products may be called experimental since they were not commercially successful. The *Atlas of Science* is emblematic of ISI in several ways. It was a cutting-edge venture rooted in a desire to understand the nature and potential of citation data, and it was a product offered without knowing whether there was any market for it. But Garfield valued the results of Small's research and took a risk.

Small synthesized the interests and perspectives of Price (mapping the research front), Merton (revealing socio-cognitive relationships and structures), and Garfield (exploiting the full richness of citation data), and he then extended these by describing the symbolic function of highly cited papers, exploring changing structures over time including identification of emerging topics, probing the context of citations, and more (Small, 2003; Pendlebury, 2013).

Citation Data in Research Evaluation

The use of citation data in the assessment of research performance relies heavily on Merton's description of the normative behavior of scientists as represented in the literature: the papers and people accruing many "credits" do so because the research community has depended on them more than others, signifying influence, impact, significance, utility, and other such notions (Garfield and Welljams-Dorof, 1992; Moed, 2005; Moed, 2017; Aksnes et al., 2019)

With the appearance of NSF's *Science Indicators* (1973), Francis Narin's *Evaluative Bibliometrics* (Narin, 1976), and *Toward a Metric of Science* (Elkana et al., 1978), interest in and use of publication and citation data as research performance indicators gathered pace. The launch of the journal *Scientometrics* (1978), initiated by Tibor Braun (1932-) of the Hungarian Academy of Sciences, also signaled that there was now a critical mass of scholars focusing on quantitative studies of science (and drawing on *SCI* data for their research). In the 1980s Braun's Information Science and Scientometrics Unit (ISSRU) focused on studies comparing the performance of nations. The ISSRU team demonstrated the use of relative citation impact and other measures. Meanwhile, the Centre for Science and Technology Studies (CWTS) at Leiden University, under the direction of Anthony F.J. van Raan (1945-), explored applications of publication and citation data to gauge performance at the meso level, that is, of universities and research groups within them. Both groups used ISI's data, but they modified and edited the data for the purposes of scientometric studies, since the data were indexed and organized for the different purpose of information retrieval.

Sociologist Jonathan Cole (1942-) of Columbia University, another student of Merton, noted that "the creation of the *SCI* represents a good case study of how technological innovations very frequently create the necessary conditions for advance in scientific fields" (Cole, 2000). In this case, it gave birth to the field of scientometrics.

Meanwhile at ISI, its research department, under the direction of Small, launched the newsletter *Science Watch: Tracking Performance and Trends in Basic Research* in 1990. The monthly publication included a feature story based on publication and citation data, an interview with a top-cited scientist, and top-ten lists of hot papers in biology, clinical medicine, chemistry, and physics, which resembled music pop-charts or bestseller book lists. During the 1990s at ISI, the research group began to produce products, first issued on diskette and then CD-ROM, for PC use: *National Science Indicators*, *University Science Indicators*, *Journal Performance Indicators*, *Highly Cited Papers*, *Local Journal Utilization Report* (used for journal collection development), and others including custom versions. In 2001, under the ownership of Thomson Corporation, Small and colleagues developed *Essential Science Indicators* (*ESI*), a web product that provides data on most-cited authors, institutions, nations, and journals in 22 broad fields, as well as highly cited and hot papers. Small's research front data were also included, with the fronts derived from co-citation analysis of ESI's highly cited papers. In 2009, under ownership of Thomson Reuters, *InCites*, a web platform for research performance analysis, was introduced.

Fueling the activity of the ISSRU, CWTS, ISI, and other groups in the 1990s was demand from government agencies and universities responding to a new regime of accountability and increasing competition for research funding. New Public Management, in Australia, the UK, and US, dramatically changed the administration of universities, moving control from the faculty senate to professional administrators who employed business models in running their institutions. Publication output and citation impact data supplied the need for so-called hard evidence about research performance and demonstrations of "value for money."

Garfield watched the sea change in use of citation data and in 1998 asked, "Is the tail now wagging the dog?" (Garfield, 1998). The dog was citation data for information retrieval and the tail scientometric data on research performance. Garfield's intellectual interests in use of his own data were, in order: information retrieval, the history and sociology of science including science mapping, and, last, research evaluation.

Garfield frequently wrote about most-cited scientists and pointed out that citation data could balance instances of biased peer review, adding fairness to an assessment. But mostly he considered the use of citation data in the appointment, promotion, and funding of researchers complex and, with Merton, "ancillary to detailed judgments by informed peers" (Merton, 1979). While he did not proscribe the use of citation data for evaluating individuals, Garfield realized that when used at scale by non-experts misuse was inevitable and would have profound career and personal consequences. As early as 1963, Garfield condemned "promiscuous and careless use of quantitative citation data for sociological evaluations, including personnel and fellowship selection" (Garfield, 1963; also Garfield, 1983).

At a presentation in Dalian, China in 2009, Garfield said the tail of scientometrics had grown into a "monster." Developments contributing to this expansion in the last two decades, among others, have been the proliferation of university rankings following the appearance of the *Academic Ranking of World Universities* in 2003, the introduction of *Google Scholar* in 2004, and the arrival of the h-index in 2005.

Researchers themselves, who in past decades opposed evaluation by citations measures, increasingly display and even tout their total citation count or h-index, a sign of the institutionalization of "metrics." In some evaluative contexts, rewards are so outsized that researchers, albeit a minority, engage in strategic behavior in publishing and citing, effectively destroying the meaning and value of the performance indicators (de Rijcke et al. 2016). On the other hand, the robust development of science mapping during the last two decades is a salutary counterweight to runaway research assessment schemes and citation gaming, as well as a return to original notions of Garfield and of Small on the meaning, richness, and value of the data.

After ISI

After the sale of ISI in 1992, Garfield served as Chairman Emeritus to the company he founded. He continued a program of personal research and publishing and frequently traveled to lecture and receive honors.

A lifelong interest in communicating scientific information within the profession inspired Garfield in 1986 to create *The Scientist*, a newspaper for researchers. When ISI was sold to Thomson Corporation, he retained *The Scientist* and developed it further in partnership with publisher Vitek Tracz (1940-) after 2002 (Grimwade, 2018). It continues today under different ownership.

In addition to his work with *The Scientist*, Garfield returned to research begun in 1964, when he mapped direct citation links between key papers on the structure of DNA research, ordered chronologically to show descendants and antecedents (Garfield et al., 1964). In the first decade of the twenty-first century Garfield and colleagues developed the software tool *HistCite*. This PC program analyzed a collection of papers imported from *WoS* and drew a map of direct citations between papers arranged by year. He called this type of analysis algorithmic historiography (Garfield et al., 2003; Garfield, 2009; Grimwade, 2018). *HistCite* has now been superseded by *CitNetExplorer*, whose developers acknowledged inspiration from *HistCite* (van Eck and Waltman, 2014).

Accolades

In recognition of his research contributions, Garfield received several honorary doctoral degrees and numerous awards, including the first Derek J. de Solla Price Memorial Medal of the journal *Scientometrics* (1984). He was a Fellow of the American Association for the Advancement of Science (1966), the American Academy of Arts and Sciences (2005), and a Member of the American Philosophical Society (2007). During 1998 – 2000, he served as President of the American Society for Information Science & Technology.

In 2000, Garfield received a festschrift volume from colleagues entitled *The Web of Knowledge* (Cronin and Atkins, 2000). He died in February 2017, at the age of 91 (Clarivate Analytics, 2017; Small. 2017; Wouters, 2017).

Transcriptions of two in-depth interviews and an extensive video interview conducted by Small provide further information on his life and career (Garfield, 1987; Garfield, 1997; Garfield, n.d.). The Science History Institute, Philadelphia, maintains an archive of Garfield's professional papers.

As mentioned, the Thomson Corporation acquired ISI in 1992. Thomson merged with Reuters in 2008 to become Thomson Reuters. In 2016, Thomson Reuters sold its intellectual property and science business, including the substance of the original ISI, to two private equity firms in partnership, Onex Corporation and Baring Private Equity Asia. The company was then rebranded as Clarivate Analytics. The ISI identity was revived at the beginning of 2018 to designate a new research division within Clarivate.

References

Aksnes, DW, Langfeldt, L & Wouters, P 2019, 'Citations, Citation Indicators, and Research Quality: An Overview of Basic Concepts and Theories', *Sage Open*, vol. 9, no. 1, article number 2158244019829575, https://journals.sagepub.com/doi/full/10.1177/2158244019829575 (July 15, 2020), doi: https://doi.org/10.1177/2158244019829575.

Archambault, E & Larivière, V 2009, 'History of the Journal Impact Factor: Contingencies and Consequences', *Scientometrics*, vol. 79, no. 3, pp. 639 – 653, doi: https://doi.org/10.1007/s11192 – 007 – 2036-x.

Bensman, SJ 2007, 'Garfield and the Impact Factor', *Annual Review of Information Science and Technology*, vol. 41, pp. 93 – 155, http://garfield.library.upenn.edu/bensman/bensmanegif2007.pdf (July 15, 2020), doi: https://doi.org/10.1002/aris.2007.1440410110.

Cawkell, T & Garfield, E 2001, 'Institute for Scientific Information', in EH Fredricksson (ed.), *A Century of Scientific Publishing*, pp. 149 – 160, IOS Press, Amsterdam, http://www.garfield.library.upenn.edu/papers/isichapter15centuryofscipub149 - 160y2001.pdf (July 15, 2020), ISBN 13: 9781586031480.

Clarivate Analytics, *Commemoration and Celebration of the Life of Eugene Garfield, 1925 – 2017*, program for event, September 15 – 16, 2017, Clarivate Analytics, Philadelphia, PA.

Cole, JR 2000, 'A Short History of the Use of Citations as a Measure of the Impact of Scientific and Scholarly Research', in B Cronin & HB Atkins (eds.), *The Web of Knowledge: A Festschrift* *in Honor of Eugene Garfield*, pp. 281–300, Information Today, Inc., Medford, NJ, http://www.garfield.library.upenn.edu/webofknowledge.html (July 15, 2020), ISBN-13: 978–1573870993.

Cronin, B & Atkins, HB (eds.) 2000, *The Web of Knowledge: A Festschrift in Honor of Eugene Garfield*, Information Today, Medford, NJ, http://www.garfield.library.upenn.edu/webofknowledge.html (July 15, 2020), ISBN-13: 978–1573870993.

de Araújo, PC, Castanha, RCG & Hjørland, B 2019, 'Citation Indexing and Indexes', in B Hjørland & C Gnoli (eds.), *ISKO Encyclopedia of Knowledge Organization*, ISKO, https://www.isko.org/cyclo/citation (July 15, 2020).

De Bellis, N 2009, *Bibliometrics and Citation Analysis: From the Science Citation Index to Cybermetrics*, The Scarecrow Press, Inc., Lanham, MD, ISBN-13: 978–0810867130.

de Rijcke, S, Wouters, PF, Rushforth, AD, Franssen, TP & Hammarfelt, B 2016, 'Evaluation Practices and Effects of Indicator Use – A Literature Review', *Research Evaluation*, vol. 25, no. 2, pp. 161–169, doi: https://doi.org/10.1093/reseval/rvv038.

Elkana, Y, Lederberg, J, Merton, RJ, Thackray, A & Zuckerman, H 1978, *Toward a Metric of Science: The Advent of Science Indicators*, John Wiley & Sons, New York, ISBN: 471 98435–3.

Garfield, E 1955, 'Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas', *Science*, vol. 122, no. 3159, pp. 108–11, http://garfield.library.upenn.edu/papers/science1955.pdf (July 15, 2020), doi: https://doi.org/10.1126/science.122.3159.108.

Garfield, E & Sher, IH 1963, 'New Factors in the Evaluation of Scientific Literature Through Citation Indexing', *American Documentation*, vol. 14, no. 3, pp. 195–201, http://www.garfield.library.upenn.edu/essays/v6p492y1983.pdf (July 15, 2020), doi: https://doi.org/10.1002/asi.5090140304.

Garfield, E 1963, 'Citation Indexes in Sociological and Historical Research', *American Documentation*, vol. 14, no. 4, pp. 289–91, http://www.garfield.library.upenn.edu/essays/V1p043y1962-73.pdf (July 15, 2020), doi: https://doi.org/10.1002/asi.5090140405.

Garfield, E 1964, '*Science Citation Index* – A New Dimension in Indexing', *Science*, vol. 144, no. 3619, pp. 649–654, http://www.garfield.library.upenn.edu/essays/v7p525y1984.pdf, doi: https://doi.org/10.1126/science.144.3619.649 (July 15, 2020).

Garfield, E, Sher, IH & Torpie, RJ 1964, 'The Use of Citation Data in Writing the History of Science', Institute for Scientific Information, Philadelphia, PA, http://www.garfield.library.upenn.edu/papers/useofcitdatawritinghistofsci.pdf (July 15, 2020).

Garfield, E & Sher, IH 1967, 'ASCA (Automatic Subject Citation Alert): A New Personalized Current Awareness Service for Scientists', *American Behavioral Scientist*, vol. 10, no. 5, pp. 29–32, http://www.garfield.library.upenn.edu/essays/v6p514y1983.pdf (July 15, 2020), doi: https://doi.org/10.1177/000276426701000507.

Garfield, E 1972, 'Citation Analysis as a Tool in Journal Evaluation', *Science*, vol. 178, no. 4060, pp. 471–479, http://www.garfield.library.upenn.edu/essays/V1p527y1962-73.pdf (July 15, 2020), doi: https://doi.org/10.1126/science.178.4060.471.

Garfield, E 1976, 'Significant Journals of Science,' *Nature*, vol. 264, no. 5587, pp. 609–617, http://www.garfield.library.upenn.edu/essays/v3p130y1977-78.pdf (July 15, 2020), doi: https://doi.org/10.1038/264609a0.

Garfield, E 1993, *Essays of an Information Scientist. Vols. 1–15*, ISI Press, Philadelphia, PA, http://www.garfield.library.upenn.edu/essays.html (July 15, 2020).

Garfield, E 1984, 'How to Use Citation Data for Faculty Evaluation, and When is it Relevant? Part 1', *Current Contents*, vol. 44 (October 31, 1983), reprinted in E Garfield, *Essays of an Information Scientist. Vol 6 (1983)*, pp. 354–362, ISI Press, Philadelphia, PA, http://www.garfield.library.upenn.edu/essays.html (July 15, 2020).

Garfield, E 1987, interview by Arnold Thackray and Jeffrey L. Sturchio, Beckman Center for the History of Chemistry, November 16, http://www.garfield.library.upenn.edu/oralhistory/interview.html (July 15, 2020).

Garfield, E 1979, *Citation Indexing – Its Theory and Application in Science, Technology, and Humanities*, http://www.garfield.library.upenn.edu/ci/title.pdf (July 15, 2020), John Wiley & Sons, New York, ISBN-13: 978 – 0471025597.

Garfield, E & Welljams-Dorof, A 1992, 'Citation Data: Their Use as Quantitative Indicators for Science and Technology Evaluation and Policymaking', *Science and Public Policy*, vol. 19, no. 5, pp. 321 – 327, http://www.garfield.library.upenn.edu/papers/sciandpubpolv19%285%29p321y1992.html (July 15, 2020), doi: https://doi.org/10.1093/spp/19.5.321.

Garfield, E 1997, interview by RV Williams, Chemical Heritage Foundation, July 29, http://garfield.library.upenn.edu/papers/oralhistorybywilliams.pdf (July 15, 2020).

Garfield, E 1998, 'From Citation Indexes to Informetrics: Is the Tail Now Wagging the Dog?', *Libri*, vol. 48, no. 2, pp. 67 – 80, http://www.garfield.library.upenn.edu/papers/libriv48%282%29p67-80y1998.pdf (July 15, 2020), doi: https://doi.org/10.1515/libr.1998.48.2.67.

Garfield, E 2000a, 'Recollections of Irving H. Sher 1924 – 1996: Polymath/Information Scientist Extraordinaire', *Journal of the American Society for Information Science and Technology*, vol. 52, no. 14, pp. 1197 – 1202, http://garfield.library.upenn.edu/papers/sherjasis&t52%2814%29p1197y2001.pdf (July 15, 2020), doi: https://doi.org/10.1002/asi.1187.

Garfield, E 2000b, 'From Laboratory to Information Explosions. The Evolution of Chemical Information Services at ISI', *Journal of Information Science*, vol. 27, no. 2, pp. 119 – 125, http://www.garfield.library.upenn.edu/papers/jis27%282%29p119y2001.pdf (July 15, 2020), doi: https://doi.org/10.1177/0165551014233626.

Garfield, E, Pudovkin AI &. Istomin, VS 2003, 'Why Do We Need Algorithmic Historiography?', *Journal of the American Society for Information Science and Technology*, vol. 54, no. 5, pp. 400 – 412, http://garfield.library.upenn.edu/papers/jasist54%285%29400y2003.pdf (July 15, 2020), doi: https://doi.org/10.1002/asi.10226.

Garfield, E 2006, 'The History and Meaning of the Journal Impact Factor,' *JAMA – Journal of the American Medical Association*, vol. 293, no. 1, pp. 90 – 93, http://garfield.library.upenn.edu/papers/jamajif2006.pdf (July 15, 2020), doi: https://doi.org/10.1001/jama.295.1.90.

Garfield, E 2009, 'From the Science of Science to Scientometrics: Visualizing the History of Science with *HistCite* Software', *Journal of Informetrics*, vol. 3, no. 3, pp. 173 – 179, http://garfield.library.upenn.edu/papers/issispain2007.pdf (July 15, 2020), doi: https://doi.org/10.1016/j.joi.2009.03.009.

Garfield, E, personal webpages, http://www.garfield.library.upenn.edu/ (July 15, 2020).

Garfield, E n.d., video interview, Web of Stories, https://www.webofstories.com/play/eugene.garfield/1 (July 15, 2020).

Grimwade, A 2018, 'Eugene Garfield – 60 Years of Invention and Innovation', *Frontiers in Research Metrics and Analytics*, vol. 3, no. 14, https://www.frontiersin.org/articles/10.3389/frma.2018.00014/full (July 15, 2020), doi: https://doi.org/10.3389/frma.2018.00014.

Lawlor, B 2014, 'The Institute for Scientific Information: A Brief History', in LR McEwen & RE Buntrock (eds.), *The Future of the History of Chemical Information*, ACS Symposium Series 1164, pp. 109 – 136, American Chemical Society, Washington, ISBN-13: 978 – 0841229457.

Larivière, V & Sugimoto, CR 2019, 'The Journal Impact Factor: A Brief History, Critique, and Discussion of Adverse Effects', in W. Glänzel, HF Moed, U Schmoch & M Thelwall (eds.), *Springer Handbook of Science and Technology Indicators*, pp. 3 – 24, Cham, Switzerland, Springer Nature, ISBN-13: 978 – 3030025106.

Lazerow, S 1974, 'Institute for Scientific Information', in A Kent, H Lancour & JE Daily (eds.), *Encyclopedia of Library and Information Science*, vol. 12, pp. 89 – 97, Marcel Dekker, New

York, http://www.garfield.library.upenn.edu/essays/v2p197y1974-76.pdf (July 15, 2020), ISBN-13: 9780824720124.

Lederberg, J 1963, 'Preface', in E Garfield & IH Sher, *Genetics Citation Index*, Institute for Scientific Information, Philadelphia, PA, http://garfield.library.upenn.edu/essays/v2p189y1974–76.pdf (July 15, 2020).

McVeigh, ME 2017, 'Citation Indexes and the Web of Science', in JD McDonald & M Levine Clark (eds.), *Encyclopedia of Library and Information Sciences*, fourth edition, vol. 2, 940–950, CRC Press, Boca Raton, FL, ISBN-13: 978–1466552593.

Merton, RK 1979, 'Foreword', in E Garfield, *Citation Indexing – Its Theory and Application in Science, Technology, and Humanities*, pp. xiii-xv, John Wiley & Sons, New York, http://garfield.library.upenn.edu/ci/foreword.pdf (July 15, 2020), ISBN-13: 978–0471025597.

Merton, RK 1988, 'The Matthew Effect in Science, II: Cumulative Advantage and the Symbolism of Intellectual Property', *ISIS*, vol. 79, no. 4, pp. 606–623, http://garfield.library.upenn.edu/merton/matthewii.pdf (July 15, 2020), doi: https://doi.org/10.1086/354848.

Moed, HF 2005, *Citation Analysis in Research Evaluation*, Springer Nature, Dordrecht, Netherlands, ISBN-13: 978–1402037139.

Moed, HF 2017, *Applied Evaluative Informetrics*, Springer Nature, Cham, Switzerland, ISBN-13: 978–3319605210.

Narin, F 1976, *Evaluative Bibliometrics: The Use of Publication and Citation Data in the Evaluation of Scientific Activity*, Computer Horizons Inc., Cherry Hill, NJ.

Pendlebury, DA & Adams, J 2012, 'Comments on a Critique of the Thomson Reuters Journal Impact Factor', *Scientometrics*, vol. 92, no. 2, pp. 395–401, doi: https://doi.org/10.1007/s11192-012-0689–6.

Pendlebury, DA 2013, 'Research Fronts: In Search of the Structure of Science', in C King & DA Pendlebury, *Research Fronts 2013: 100 Top-Ranked Specialties in the Sciences and Social Sciences*, pp. 29–31, Thomson Reuters Philadelphia, PA, http://garfield.library.upenn.edu/papers/pendleburykingresearchfronts2013.pdf (July 15, 2020).

Price, DJ de Solla 1961, *Science Since Babylon*, Yale University Press, New Haven, CT, ISBN 13: 9780300017984.

Price, DJ de Solla 1963, *Little Science, Big Science*, Columbia University Press, New York, NY, ISBN-13: 978–0231085625

Price, DJ de Solla 1965, 'Networks of Scientific Papers', *Science*, vol. 149, no. 3683, pp. 510–515, http://garfield.library.upenn.edu/papers/pricenetworks1965.pdf (July 15, 2020), doi: https://doi.org/10.1126/science.149.3683.510.

Price, DJ de Solla 1976, 'General Theory of Bibliometric and other Cumulative Advantage Processes', *Journal of the American Society for Information Science*, vol. 27, nos. 5–6, pp. 292–306, http://garfield.library.upenn.edu/price/pricetheory1976.pdf (July 15, 2020), doi: https://doi.org/10.1002/asi.4630270505.

Price, DJ de Solla 1980, 'Foreword', in E Garfield, *Essays of an Information Scientist. Volume 3. 1977–1978*, pp. v-ix, ISI Press, Philadelphia, PA.

Schnell, JD 2018, 'Web of Science: The First Citation Index for Data Analytics and Scientometrics', in FJ Cantú-Ortiz (ed.), *Research Analytics: Boosting University Productivity and Competitiveness through Scientometrics*, pp. 15–29, CRC Press, Boca Raton, FL, ISBN-13: 978–1498785426.

Small, H 1973, 'Co-citation in the Scientific Literature: A New Measure of the Relationship between Two Documents', *Journal of the American Society for Information Science*, vol. 24, no. 4, pp. 265–269, http://www.garfield.library.upenn.edu/essays/v2p028y1974-76.pdf (July 15, 2020), doi: https://doi.org/10.1002/asi.4630240406.

Small, H & Garfield, E 1985, 'The Geography of Science: Disciplinary and National Mappings,' *Journal of Information Science*, vol. 11, no. 4, pp. 147–159, http://www.garfield.library.upenn.

edu/essays/v9p325y1986.pdf (July 15, 2020), doi: https://doi.org/10.1177/016555158501100402.

Small, H 2003, 'Paradigms, Citations, and Maps of Science: A Personal History', *Journal of the American Society for Information Science and Technology*, vol. 54, no. 5, pp. 394–399, doi: https://doi.org/10.1002/asi.10225.

Small, H 2017, 'A Tribute to Eugene Garfield: Information Innovator and Idealist', *Journal of Informetrics*, vol. 11, no. 3, pp. 599–612, doi: https://doi.org/10.1016/j.joi.2017.04.006.

van Eck, NJ, & Waltman, L 2014, 'CitNetExplorer: A New Software Tool for Analyzing and Visualizing Citation Networks', *Journal of Informetrics*, vol. 8, no. 4, pp. 802–823, doi: https://doi.org/10.1016/j.joi.2014.07.006, see https://www.citnetexplorer.nl/ (July 15, 2020).

Weinstock, M 1971, 'Citation Indexes', in A Kent (ed.), *Encyclopedia of Library and Information Science*, vol. 5, 16–40, Marcel Dekker, New York, http://www.garfield.library.upenn.edu/essays/V1p188y1962-73.pdf (July 15, 2020), ISBN 13: 9780824721053.

Wouters, P 1999, *The Citation Culture*, PhD thesis, University of Amsterdam, Amsterdam, http://garfield.library.upenn.edu/wouters/wouters.pdf (July 15, 2020).

Wouters, P 2017, 'Eugene Garfield (1925–2017): Inventor of the *Science Citation Index*', *Nature*, vol. 543, no. 7646, p. 492, https://www.nature.com/articles/543492a (July 15, 2020), doi: https://doi.org/10.1038/543492a.

Zuckerman, H 2018, 'The Sociology of Science and the Garfield Effect: Happy Accidents, Unanticipated Developments and Unexploited Potentials', *Frontiers in Research Metrics and Analytics*, vol, 3, no. 20, https://www.frontiersin.org/articles/10.3389/frma.2018.00020/full (July 15, 2020), doi: https://doi.org/10.3389/frma.2018.00020.

1.4 Derek De Solla Price: The Father of Scientometrics

Frashid Danesh and Ali Mardani-Nejad

Abstract: The 1960s was a golden age in the quantitative sciences. The decade we witnessed was the age of different disciplines and studies expansion in science. One of the great pioneers and avant-gardes in the development of these fields can be called Derek De Solla Price. Price's descriptions of "science of science" have led to the definition of scientometrics and he is famous as the father of this field. Price was one of those first who introduced "x number of authors" as one of the indicators of scientific activities in different countries. The costly activities of scientometrics were so valuable that Tibor Brown, the founder and editor-in-chief of the journal *Scientometrics*, created a prize in 1983 to commemorate Price's work. This prize is the first and foremost international infometric prize awarded every two years. This chapter will cover his biography and his activities, and topics such as science of science, citation and referencing, the Price index, and his scientific growth pattern.

Keywords: understanding De Sola Price, scientometrics, Price index, scientific growth pattern, biology, science of science, citation and referencing.

Derek J. de Solla Price Biography

Derek John Price was born on January 22, 1922, in Leyton, a suburb of London, England, to Fanny Marie de Solla, a singer, and Philip Price, a tailor. Both of his parents were descended from Jewish immigrant families. He spent the 1946 – 1947 academic year at Princeton University as a Commonwealth Fund Fellow in mathematical physics and married Ellen Hjorth of Copenhagen, Denmark, in 1947. They had two sons and a daughter. Derek Price died in September 1983, after suffering a massive heart attack.

Introduction

After presenting a description of De Sola Price biography and a glimpse at a timeline of his activities during his lifetime, there will be some explanation of his fundamentals, thoughts, and important works in this chapter. Discussion of topics will include

Farshid Danesh, Assistant Professor, Information Management Research Department, Regional Information Center for Science & Technology (RICeST), Shiraz, Iran, farshiddanesh@ricest.ac.ir (Corresponding Author)
Ali Mardani-Nejad, Young Researchers and Elite Club, Najafabad Branch, Islamic Azad University, Najafabad, Iran

https://doi.org/10.1515/9783110646610-006

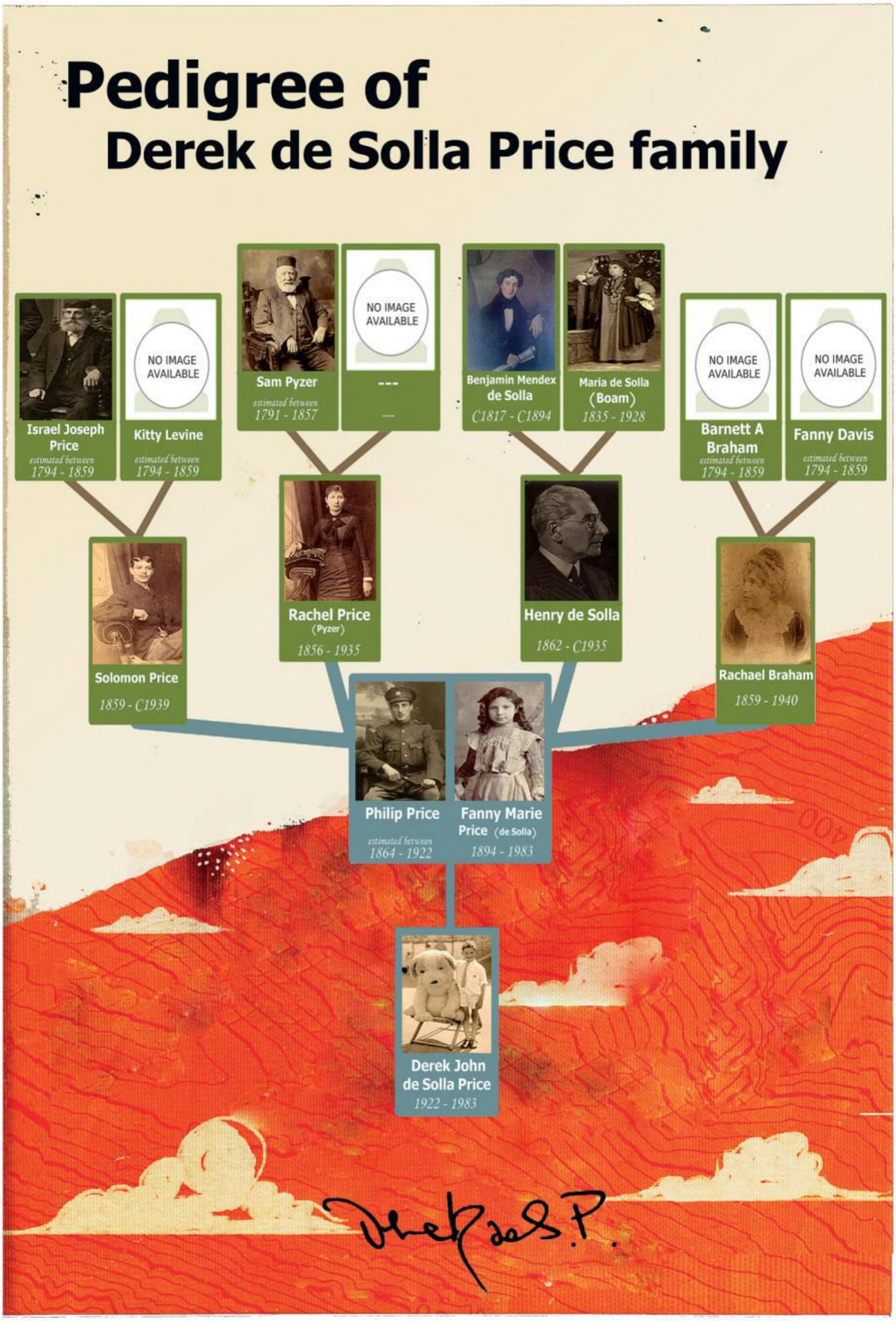

Figure 1: Pedigree of Derek de Solla Price family (MyHeritage Company, 2019).

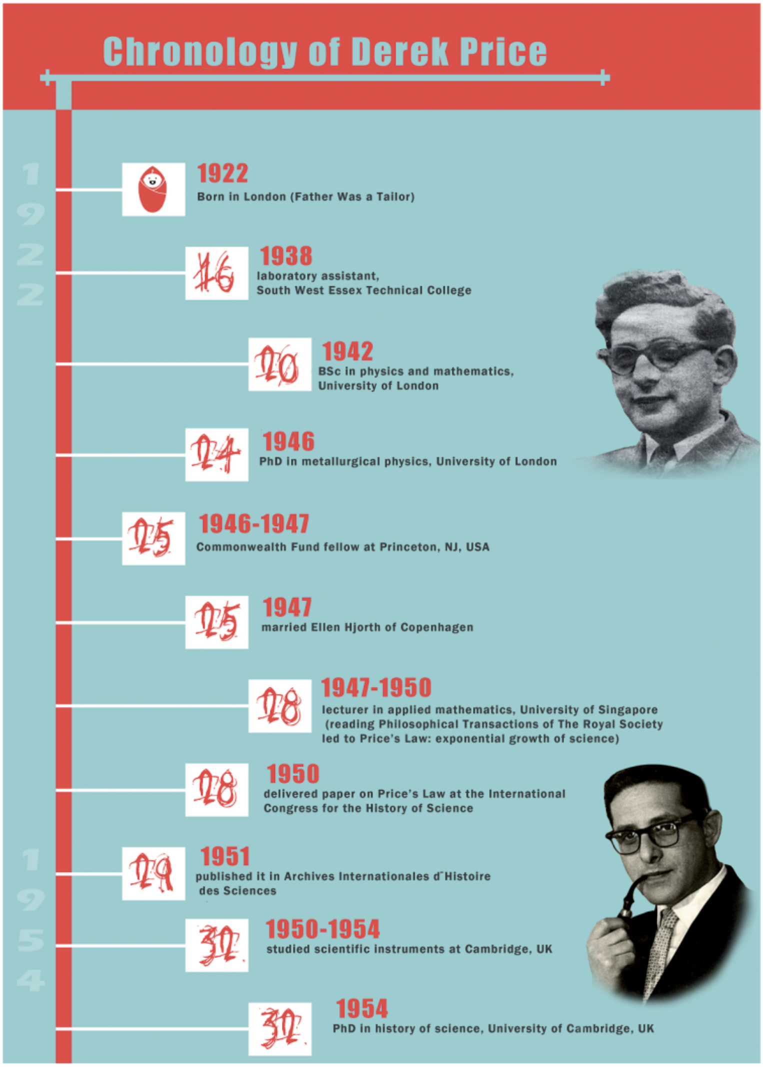

Figure 2: Chronology of Derek De Solla Price's life and profession (Alchetron Free Social Encyclopedia, 2018; McGill University, 1976).

science of science, citation and referencing, scientific growth pattern, and, at the end, the Price index, which he has a very colorful role in. During this chapter, these issues will indicate the secret of his immortality in scientometrics.

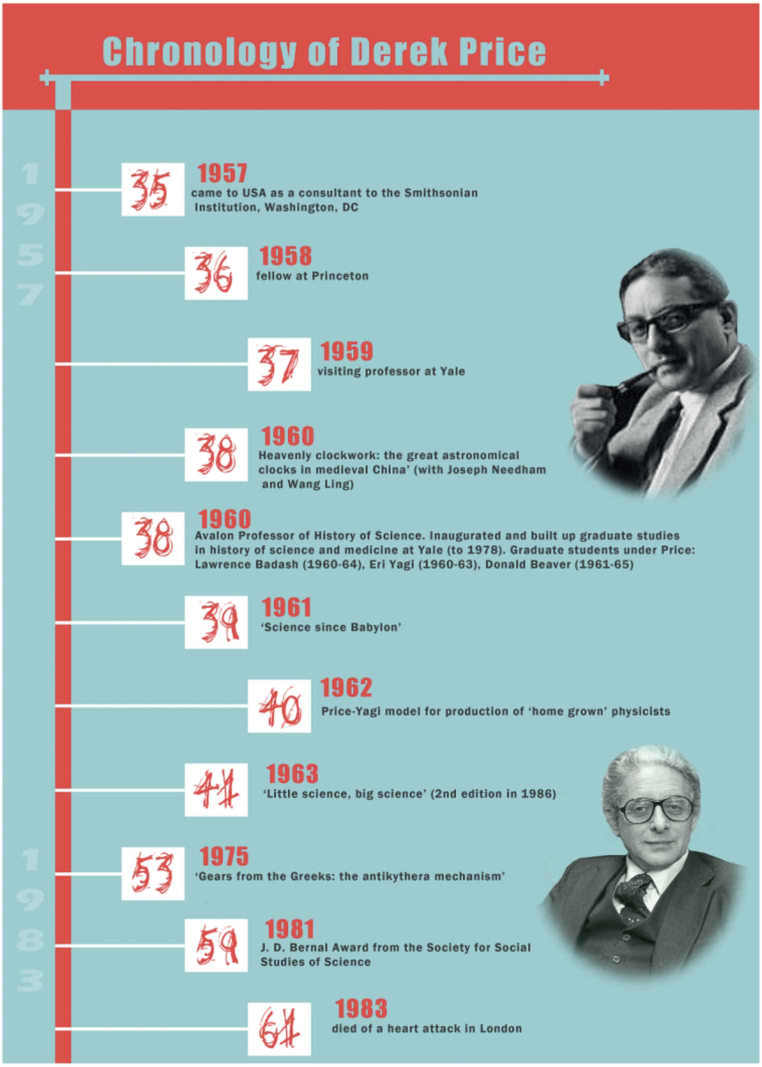

Figure 2 (continued): Chronology of Derek De Solla Price's life and profession (Alchetron Free Social Encyclopedia, 2018; McGill University, 1976).

Science of Science – Scientometrics

Price is one of the founders of science of science, and is one of the prominent representatives of the physical approach in quantitative studies of science and scientific activity as well as the creation of scientometrics indicators. A few years before the Russian scientist Vasiliy Vasilievich Nalimov (Nakometria, 1969) invented scientometrics, which led to the creation of a journal of the same name (1978), the description of the science of science led to the definition of scientometrics. Price's definition of the science of science is that the various scientific indicators must link with simple laws, in such a way that their phenomenological interpretation is possible.

The first chapter in Price's 1963 book, *Little Science, Big Science*, is called "A Science of Science." The fourth sample reference quotes from the Preface:

Price, on morphology of the term "scientometrics", in the third sample reference, comments on some of this etymology. He is quoted from the enlarged 1975 edition of a book originally published in 1961:

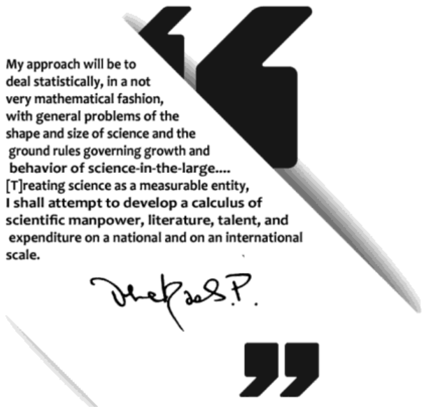

Figure 3: Price, 1963, p. viii quoted in Diodato 1994, p. 147.

The material covered in this chapter has probably undergone more development and change than any other. It rapidly proved to have a life of its own, so that it grew first into a separate book (*Little Science, Big Science*) and then touched off a continuing series of research papers exploring many different quantitative investigations based on the counting of journals, papers, authors, and citations. In no time at all there were bibliographies and conventions devoted to bibliometrics and scientometrics, and even a meeting of the invisible college or people studying invisible colleges. ... [T]he term "science of science" achieved an almost explosive popularity. Unfortunately, though it came readily to the tongue and pleased those who desired objective investigations of the workings of science in society, the term rapidly became debased by being used in as many different ways, as there were users, and by being taken as a promise to deliver goods that were by their very nature undeliverable. (Price, 1963, p. viii, quoted in Diodato, 1994, p. 146)

Derek de Solla Price asserted that the primary aim of scientometrics is to do scientific analysis of science mathematically. He lucidly described that however successful we are in understanding the productivity of scientists and the mechanics and pattern of coherence of scientific creativity, we shall still need historians, sociologists, and psychologists of science for those types of analyses that cannot be expressed in metric terms (Sengupta, 1992, p. 86). Elsewhere, he asserted that: "Somewhat cautiously, it may be suggested that we need a social scientific equivalent of the

Newtonian masterstroke that took such vaguely used terms as force, work, and energy, redefined them with simple equations [...] and brought order into previous meanderings" (Price, 1980b, p. 1, quoted in Moed, 2005, p. 195).

Citation and Referencing

Derek De Solla Price stated that citation and referencing to earlier works in the West have been used since 1850 in scientific journals. The evidence suggested that the predecessors were aware of the necessity of using citation and referencing, and in their works they cited past works (Haghighi, 2003).

Derek De Solla Price studied how to conduct scientometrics and a network of citations among academic papers. He indicated that articles cited many times are likely to be cited in the future. The opposite is also true of this phenomenon; articles not cited regularly will not likely be cited in the future, which led to a model for this phenomenon. He also studied scientometrics to evaluate research. On this basis, Derek De Solla Price opened the door to a valuable contemporary bibliometric (Glanzel, 2008): "It seems to me a great pity to waste an excellent technical term by using the words citation and reference interchangeably. I therefore propose and adopt the convention that if Paper R contains a bibliographic footnote using and describing Paper C, then R contains a reference to C, and C has a citation from R. The number of references a paper has measured by the number of items in its bibliography as endnotes and footnotes, etc., while the number of citations a paper has is found by looking it up on some sort of citation index and seeing how many other papers mention it" (Price, 1970, p. 3, quoted in Moed, 2005, p. 114).

Price's Model

In 1963, Derek De Solla Price offered a model for explaining the productivity of authors in a thematic field. After analyzing the ideas of Francis Galton in the characterization of the elite and Alfred Lotka in productivity in chemistry and physics, he proposed a particular model (Sotudeh and Yaghtin, 2014).

Price's Law

This describes the number of prolific authors in a subject field. In a given field during a given period, the number of prolific authors is equal to approximately the square root of the total number of authors in the field. In particular, the prolific authors account for about half the publications in the field. This was named for Derek J. de Solla Price, and is also called Price's square root law (Diodato 1994, p. 131). Rousseau's law is one of the sources for Price's law of authorship, which in turn also draws ideas from Lotka's law of authorship. The sample references below suggest that the naming of the law may be more of attribution of an idea rather than a claim that Rousseau discussed square roots (Diodato 1994, p. 138).

Price's law, as said also known as the Price's Square Root Law, focuses on the relationship between the literature and the number of authors in the subject area. He stated that half of the publications come from the square root of all contributors. Thus, if 25 authors write 100 papers, five authors will have contributed 50 papers (Nicholls, 1988). Derek de Solla Price predicted that the number of elites in science is small compared to the total number of scientists. In his law, he claimed that any population of size N contains a capable elite of size /N. Alternatively, in other words, "One-half of all scientific papers contributed by several authors equal to the square root of the total number of scientific authors" (Sengupta, 1992, p. 80).

In the first sample reference, Price (1963) derived the law after discussing the ideas of Francis Galton (on elitism) and Alfred Lotka (on authorship in chemistry and physics). Price said: "If one computes the total production of those who write in papers, it emerges that a large number of low producers account for about as much of the total as the small number of large producers. In a simple schematic case, symmetry may be shown to a point corresponding to the square root of the total number of men, or the score of the highest producer" (Price, 1963, p. 46, quoted in Diodato 1994, 131).

Text Growth Rate

Price (1963), in the *Little Science, Big Science*, a fundamental work in the philosophy of science and scientometrics, for the first time, using statistical and objective data, depicted the phenomenon of Text Growth Rate. He showed that the number of scientific articles had doubled every 15 years between the 1660s, from scientific journal creation until the 1960s when Derek de Solla Price's book was written. That is, if in 1660, there was only one article in a specific subject, in 1977, this number would reach 3.2 million articles (Heidari, 2011).

Price suggested the total number of journals in 1980 to be about 40,000 (Price, 1980a). To paraphrase Eugene Garfield and Robert Merton, who laughed at him in a posthumous work, an expanded version of *Little Science, Big Science*, we can hardly doubt that Derek de Solla Price is, indeed, the Father of Scientometrics (Fernández Cano, Torralbo, and Vallejo, 2004).

When he taught applied mathematics at Raphael College, Derek de Solla Price developed a formula for calculating the growth of science, as well as calculating the half-life of scientific papers, which played a significant role in the formation and expansion of scientometrics (Furner, 2003). The following figure visualizes the growth of the number of scientific journals and review journals since 1665. The growth rate is such that the size of science has doubled every 10 or 15 years since the seventeenth century (Fernández Cano, Torralbo, and Vallejo, 2004)

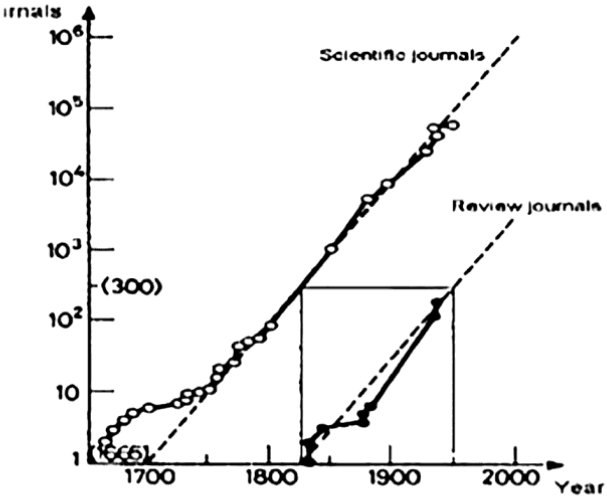

Figure 4: Total Number of Scientific Journals and Abstract Journals Founded, as a Function of Date (Price, 1963, p. 8).

Interacting with Countries with Common Scientific and Cultural Backgrounds

Using the results of scientific research, one can identify the scientific and cultural interactions of their country with other countries in recent years and the highest frequencies of such. Accordingly, recognizing the factors affecting the establishment of these communications and interactions can be necessary in order to strengthen the positive points and resolve weaknesses. In addition, this information can lead to the recognition of common areas of scientific and cultural cooperation among countries (quoted in Price, 1963; Noroozi Chakoli, 2012). It should be noted that Price was one of the first to count the number of authors as an indicator of scientific activities in different countries.

Scientific Growth Pattern

It was in 1986 when Derek de Solla Price for the first time, in his book *Little Science, Big Science*, made this discussion using scientific growth pattern. He believed that since the study of all scientific publications is not possible, it is necessary to know about science growth's rise, fall, and measurement, and he therefore suggested a study on measuring scientific growth using quantitative methods. He subsequently designed a two-variable intrinsic model that demonstrated scientific growth over time (Fernandez Cano, Torralbo and Vallejo, 2012), which gradually became more complicated with the entry of other explanatory variables, and evolved into a dynamic and vibrant promotional model.

Derek de Solla Price proposed three stages in his Scientific Growth Pattern: a) a primary stage with small increments; b) in second stage knowledge having a pure exponential growth; and c) a period of inalterable development. However, while it was argued that there is a period during which both the rates of increase and the absolute increase decline and eventually approach zero, nothing had been encountered in Derek de Solla Price's works to prove the claim (Fernández Cano, Torralbo, and Vallejo, 2004)

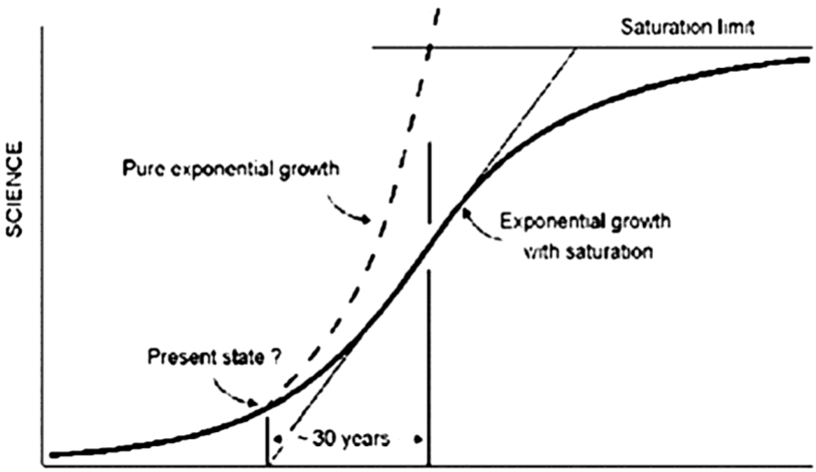

Figure 5: Price's model of scientific growth (Fernández et al., https://www.researchgate.net/profile/Antonio_Cano72004).

Finally, in an article published in 1971, Price empirically verified his complete model, using data on Ph.D. production and on scientific and manpower funds to show that there is a transition period of almost linear growth extending for two or three decades between the free exponential growth and the onset of adequate saturation (Fernández Cano, Torralbo, and Vallejo, 2004).

Price's index

According to Price (1970), this is a measure of the recency of references in a document, journal, or an entire subject field. It is the ratio of the number of references that are no more than five years old to the total number of references. It was frequently known as Price's immediacy index, but should not be confused with the measure known merely as the Immediacy Index (II); it can be compared with currency, half-life, median citation age, and recency score (Price, 1970).

The Last Word

This chapter opened using Derek de Solla Price's infographic based on his chronology, an in-depth and comprehensive look into his family life and professionalism. Derek de Solla Price was born in England in 1922 and died of a heart attack in the United States in 1983 at the age of 61. His life, in terms of length and number of years, was concise, but in terms of studying science and knowledge, he had a profound depth and various aspects.

His scientific life began with mathematics and physics and ended with scientometrics and science history. Immigration to the United States could be considered as one of the most critical events in his life; elsewhere, continuing to study at the University of Cambridge and working at Yale University, educating many students, many of whom were famous names of their era, gave Price's life, as a university professor, a great glory.

Awareness of the laws of physics allowed Derek de Solla Price to apply these rules to scientometrics. The first issue that calls his name around the world is "Science of Science." One of the things that makes him well known is his belief in scientometrics being interdisciplinary. He believed that it was possible to use the laws and principles of other sciences in scientometrics. Among Derek de Solla Price's most prominent works was *Little Science, Big Science*, as well as concepts such as "Science of Science," "progressive growth of texts," and "scientific growth pattern." Despite more than 30 years since Derek de Solla Price's death, reviewing Web of Science and Scopus suggests that his scientific production and publications are of interest to scientometrics scholars today and still feature valuable articles, books, and seminars. Peace Be Upon Him.

References

Alchetron Free Social Encyclopedia 2018, Derek J de Solla Price, https://alchetron.com/Derek-J-de-Solla-Price (Sept 22, 2020).

Diodato, V 1994, *Dictionary of bibliometrics*, The Haworth Press, New York.

Fernandez Cano, A, Torralbo, M & Vallejo, M 2012, 'Time series of scientific growth in Spanish doctoral theses (1848–2009)', *Scientometrics*, vol. 91, no. 1, pp. 16–36.

Fernandez Cano, A, Torralbo, M & Vallejo, M 2004, 'Reconsidering Price's model of scientific growth: An overview', *Scientometrics*, vol. 61, no. 3, pp. 301–321.

Furner, J 2003, 'Little book, big book: Before and after little science, big science: A review article, part II', *Journal of Librarianship and Information Science*, vol. 35, no. 3, pp. 189–201.

Glanzel, W 2008, 'De Solla Price, and the evaluation of scientometrics', *Research Trends*, September (7), pp. 4–5.

Haghighi, M 2003, 'Usage of citation in scholarly writing', *Iranian Journal of phycology and education*, vol. 32, no. 2, pp. 215–232.

Heidari, GR 2011, *Epistemology in Scientometrics*, Islamic World Science Citation Center (ISC), Shiraz.

McGill University. 1976. Articles by Beatty Lecture Archive: Derek de Solla Price, https://www.mcgill.ca/beatty/digital-archive/past-lectures/derek-de-solla-price-1976 (Sept 22, 2020).

Moed, HF 2005, *Citation analysis in research evaluation*, Springer, Netherlands.

MyHeritage company 2019, Professor, Ph.D. Derek John de Solla Price, https://www.geni.com/people/Professor-Ph-D-Derek-John-de-Solla-Price/6000000003340469413 (Sept 22, 2020).

Nalimov, VV & Mul'chenko, ZM 1969, *Naukometriya. Izuchenie nauki kak information nogo protsessa (scientometrics: Study of science as an information process)*, Nauka, Moscow.

Nicholls, PT 1988, 'Price's square root law: empirical validity and relation to Lotka's Law', *Information processing and management*, vol. 24, no. 4, pp. 469 – 477.

Noroozi Chakoli A 2012, 'The Role and Situation of the Scientometrics in Development', *Iranian Journal of Information Processing Management*, vol. 27, no. 3, pp. 723 – 736.

Price, DJD 1963, *Little Science, Big Science*, Columbia University, New York.

Price, DJD 1970, 'Citation measures of hard science, soft science, technology, and non-science', in CE Nelson & DK Pollock (eds.), *Communication among scientists and engineers*, Heath, Lexington, MA.

Price, DJD 1975, *Science since Babylon*, Yale University, New Haven.

Price, DJD 1980a, *The citation cycle*, in BC Griffith, *Key papers in information science*, pp. 195 – 210, Knowledge Industry Publications, White Plains, NY.

Price, DJD 1980b, 'Towards a Comprehensive System of Science Indicators', *Paper presented to the Conference on Evaluation in Science and Technology – Theory and Practice*, Dubrovnik, July 1980.

Sotudeh, H & Yaghtin, M 2014, 'Indicators and models for measuring researchers' scientific productivity', *Science and Technology Policy Letters*, vol. 4, no. 1, pp. 47 – 62.

1.5 Coevolution of Field and Institute: The Institutionalization of Bibliometric Research Illustrated by the Emergence and Flourishing of the CWTS

Anthony van Raan

Abstract: The quantitative study of science is mostly referred to as scientometrics. Within scientometrics, the research on scientific communication, particularly with data from publications, patents, citations, and journals, is called bibliometrics. The development of a new field of science as an "accepted" field within the academic community goes hand in hand with the rise and flourishing of institutes that shape the new field. In this chapter we illustrate this type of coevolution by a brief outline of the history and institutionalization of scientometric and particularly bibliometric research by one of the world's pioneering institutes, the Centre for Science and Technology Studies (CWTS) at Leiden University. We highlight the origin of the field, its institutionalization, the acceptance of the field within a critical academic community, the fate and fortunes of the new institute, the interplay of chance and seizing opportunities, the role of practical applications as the driving force of the development of the field, the crucial support of both public (university, ministries, research councils) and private companies, and especially of key figures who acted as champions for the promotion of the careful application of bibliometrics. The CWTS case tells of the successful synergy of pioneering work, the inspiring dynamics of bibliometrics as an internationally emerging interdisciplinary field, applications, theoretical underpinning, education, and commercialization.

Keywords: short history of scientometrics, institutionalization, CWTS, Leiden University, Elsevier, bibliometrics, research evaluation, science maps.

The quantitative study of science is mostly referred to as scientometrics. Within scientometrics, the research on scientific communication, particularly with data from publications, citations, and journals, is called bibliometrics. The development of a new field of science as an "accepted" field within the academic community goes hand in hand with the rise and flourishing of institutes that shape the new field. In this chapter we illustrate this type of coevolution by a brief outline of the history and institutionalization of scientometric and particularly bibliometric research by one of the world's pioneering institutes, the Centre for Science and Technology Studies (CWTS) at Leiden University. For comprehensive overviews of the early develop-

Dr. **Ton (A.) F.J. van Raan**, Centre for Science and Technology Studies, Leiden University, Leiden, The Netherlands, vanraan@cwts.leidenuniv.nl

https://doi.org/10.1515/9783110646610-007

ments in the quantitative studies of science we refer to Francis Narin in his seminal report *Evaluative Bibliometrics* (Narin, 1976), and to reviews of this author (van Raan, 2004; 2019).

Most of the research in bibliometric analysis was concentrated in the US until the early 1970s. This research was tremendously stimulated by the crucial breakthrough in the history of the quantitative studies of science, the creation of the *Science Citation Index* (SCI, now the Web of Science) by Eugene Garfield in his Institute for Scientific Information (ISI) (Garfield, 1955; Wouters, 1999). Cees le Pair, physicist and the then director of research at the physics research council (FOM) of the Netherlands Organization for Scientific Research (NWO), played a crucial role in the early development of bibliometric analysis outside the US. After a working visit in 1975 at the National Science Foundation and discussions with Derek de Solla Price and Eugene Garfield, le Pair immediately grasped the importance of the SCI as a support instrument for science policy by using publication and citation data in research evaluation. An important decision was made on an issue that is still current today: one of the first things we want to know is how bibliometric analysis will work in basic versus applied research. Two studies were carried out, the first in the basic research field of magnetic resonance (Chang, 1975), and the second in the applied research field of electron microscopy (Bakker, 1977).

The results of these studies clearly showed the potential as well as the limitations of bibliometric, and particularly citation, analysis. For magnetic resonance, high correlations were found for the identification of important breakthroughs based on citation analysis and based on peer assessments. For electron microscopy, however, citation analysis did not reveal all important developments and it became clear that patent analysis was indispensable in fields of applied research. In addition, towards the end of the 1970s one of the most important pioneers in bibliometrics, the Hungarian chemist Tibor Braun, established a bibliometric research group at the Academy of Sciences in Budapest. He made a crucial step for the academic emancipation of the field: the creation in 1978 of *Scientometrics*, the first journal for the quantitative research of science and technology.

How did the field of scientometrics reach Leiden University? Certainly, it was not a new university, looking for new departments. So why was this new field – often viewed with suspicion by the academic community – rapidly institutionalized in this centuries old, traditional and internationally renowned research university? What happened? In 1979 the Leiden University Executive Board decided that the allocation of the direct government funding to departments should be based, next to student numbers, on the proven quality of the research in the departments. For that time, this was a revolutionary step. I was involved in the development of this model, and as a physicist knowing Cees le Pair, I proposed using citation analysis in the assessment of research performance. Perhaps not that surprising, the research intensive Faculties of Natural Sciences and of Medicine were quite enthusiastic about the idea and in 1980 an extensive study was started: a citation analysis of all research groups in physics, astronomy, chemistry, biology, mathematics, and pharmacy, followed by all medical research groups. In total we analyzed 140 research groups, and with that it was the first large-scale bibliometric analysis at the "working floor" level. Moreover, this study also included interviews about the results with the researchers involved and, where possible, comparisons with peer review. All in all, the responses were encouraging and the Executive Board decided in 1983 to create a small group on bibliometric indicators development. The (predecessor of) CWTS was born. The report of this study (Moed et al., 1983), followed by a series of highly cited papers based on the report, catapulted the Leiden group to instant fame within the field of scientometrics. But we were not alone. I already mentioned the group of Braun in Budapest and at the University of Sussex Martin and Irvine carried out their groundbreaking bibliometric research of radio-astronomy groups in Europe (Martin and Irvine, 1983). Two more important bibliometric studies were conducted in the Netherlands around the same time: an analysis of biochemistry (Over Leven 1982) and of medical research in the Netherlands (Rigter 1983, 1986).

From 1983 onwards the brand-new Leiden group could count on continuous support of the Executive Board, the Ministry of Education, Culture and Science, and the publishing company Elsevier, mostly via contract-research projects. We also enjoyed a rapidly increasing international interest followed by increasing numbers of contract research assignments. The university provided us with a small basic financing, but otherwise we had to be self-supporting. In the beginning the SCI was only available in printed version, in very thick books, and thus, for each analysis, we had to visit the University Library. We started to use the new character-recognition facilities of the University in an attempt to transfer the printed version into a computer-readable version. This turned out to be a very cumbersome path. Fortunately, ISI soon provided us with the SCI on magnetic tapes. From that moment on we immediately started developing algorithms for data cleaning, data organization, and, the ultimate goal, to create our "enriched" SCI-based bibliometric database suitable for the accurate calculation of a broad set of bibliometric indicators.

Especially in these early days, Elsevier was particularly generous in the allocation of contract research. The Elsevier support came from the highest level: Pierre Vinken, the then CEO of Elsevier and professor of medical informatics at Leiden University was very much in favor of the development of the young bibliometric research group. In addition to typical publisher-related studies such as bibliometric analyses of journals, Elsevier gave us room to greatly broaden our bibliometric research and at the same time to go into depth, and to publish the results. This enabled us to work on increasingly advanced bibliometric methods. Particularly in the first half of the 1980s, important topics were co-citation analysis; statistics of bibliometric indicators; application of citation analysis in the social sciences and humanities; comparison of the results of bibliometric indicators with the outcomes of peer assessments; and bibliometric analysis of interdisciplinary research. Throughout this chapter we present in text boxes the topics of influential CWTS papers ranging from 1981 to 2019 in periods of five years.

Publications can be characterized by their list of references, and this forms the basis of co-citation analysis (Small, 1973). Another way to characterize publications is a list of characteristic terms, for instance specific concepts in the text of the publications, or author- and database-given keywords. This opened the way to a new development: co-word analysis, mathematically similar to co-citation analysis (Callon et al., 1983). It took quite a while, as computer power was not sufficient in the early 1980s, but finally co-word analysis became one of the basic methods to create science maps. The Leiden group played a crucial role in this development.

1981–1985
Research performance evaluation; field- and time dependence of indicators; development of the crown indicator, our flagship citation impact measure.

Initially the Leiden bibliometrics group was part of the staff of the Executive Board. However, the Board decided that given the increasing research activities the group should have a more academic basis. Sociology professor Mark van de Vall recognized the importance of quantitative research of science and technology in relation to the sociology of science and in 1986 the young group became a research unit in his Institute of Social Policy Research, Faculty of Social and Behavioral Sciences. This turned out to be a healthy basis. The new research unit was rapidly growing, albeit for about 90% "living" from contract research. It also made a further, strategically crucial next step possible: institutionalizing bibliometric research as an internationally recognized academic research field. In order to achieve this, it was crucial to set up a series of regular conferences with the highest possible standards. As discussed earlier, the field already had its own journal, *Scientometrics*, and this could be used to publish the best papers of a conference.

The Belgian mathematician Leo Egghe organized in 1987 the first international conference on bibliometrics and the theoretical aspects of information retrieval. This conference became the forerunner of today's ISSI[1] conferences. For the Leiden group this first international bibliometrics conference was the perfect place to extensively present the results of our research, and so we did. A year later we published the first *Handbook of Quantitative Studies of Science and Technology* (van Raan, 1988) and organized the first international conference specifically devoted to science and technology indicators in Leiden. From that time, this STI conference has been organized every two years, and since recently every year.

In 1989 the Leiden group acquired its formal name, *Centre for Science and Technology Studies* (CWTS), and after several organizational restructuring operations CWTS became an autonomous research department with this author as director. Important work at the end of the 1980s and early 1990s focused on bibliometric anal-

1 ISSI: International Society for Scientometrics and Informetrics, https://www.issi2019.org/ (July 15, 2020).

yses in the humanities and social sciences, the further development of co-word and co-citation analysis, and analysis of science and technology interactions on the basis of citations to the scientific literature in patents. At the same time, we further improved our standard bibliometric analyses for research performance evaluation of university departments and institutes. This work formed the firm basis of our contract research. Our applied work was much more than the use of routines: it became a source of innovation and inspiration for the continuous development of new methods and the improvement of our data analytical tools. Through this fruitful combination of applications and basic research we were able to become one of the most active and influential bibliometrics research institutes worldwide.

1986–1990
Cross-field impact; impact delay; peer review and bibliometric analysis; validation studies; quasi-correspondence analysis and multidimensional scaling of bibliometric matrices; co-word mapping; patent analysis, science and technology interface, science base of technology; co-subfield analysis; citation balances of journal relations; fractal structure of the co-citation landscape; integrating multiple sources of information.

What was the secret behind this success? Next to the competence and ambitions of our staff, I think two elements were crucial. First, our heavy investments in computers and ICT personnel. Bibliometrics is a data-intensive field that cannot function without massive computing power and development of effective data-analytical algorithms. In addition, the fact that I am a physicist may also have been instrumental in being taken seriously in a sometimes quite averse academic environment. And what has proved indispensable is that the Leiden University Board has always been well-disposed towards CWTS in terms of organizational and policy support. After a successful take-off, our task was to consolidate and further reinforce the position of CWTS.

The 1990s was a decade of numerous CWTS achievements. From the perspective of institutional academic status, the appointment of this author in 1991 to Professor of Quantitative Studies of Science, probably the first chair in this field worldwide, was crucial. It put CWTS at the level of a well-established university department, and enabled us to organize PhD work largely autonomously. In particular, the Faculties of Medicine and of Natural Sciences strongly supported the establishment of the chair. After the appointment, an annual series of lectures for MSc students was set up which has evolved into what is now the CWTS international *Graduate Course on Measuring Science.*

1991–1995
international collaboration; bibliometric analysis and interdisciplinary research; quality judgement of journals in the humanities and social sciences; co-word analysis with combined clustering and multidimensional scaling; economics research, comparison with peer review; determinants of citation scores; patent co-classification mapping

of technology; the CWTS bibliometric database; combined co-citation and word mapping; cognitive resemblance in citation relations.

In the 1990s contract research commissioned by organizations and institutions worldwide, particularly within the European Union (universities, research councils, charities), further increased. We developed standardized procedures for the execution of performance analyses in which participation of the departments or institutes is a crucial element, particularly to verify the correct assignment of publications to research groups, as well as the completeness and correctness of publications sets. Meanwhile, ISI had become part of Thomson Reuters and the SCI was migrated from CD-ROMs to internet-based facilities which eventually led to the Web of Science (WoS).

A new development was the foundation of the *Netherlands Observatory for Science and Technology* (NOWT), established in 1992 as a joint venture of CWTS and MERIT,[2] Maastricht University. Its purpose was to compile the biannual Science and Technology Indicators (WTI) Report for the Ministry of Education, Culture and Sciences. CWTS produced the WTI reports until 2010. From 1994 CWTS was involved in the VSNU national research assessment procedures.[3] For a number of major disciplines, particularly medicine, biology, chemistry, physics, and psychology, we performed extensive bibliometric analyses of all research groups. These research evaluations of many hundreds of university institutes, departments and research groups provided us with an extensive experience and expertise in the application of bibliometric research performance analysis. The presentation of our results, part of the standardized procedure, often evoked emotional reactions, particular in the sense that bibliometric analysis, especially citations, cannot capture all aspects of scientific quality. Bibliometric analysis enables the assessment of one, but certainly an important aspect of quality, international impact. I often felt like a missionary, preaching with full devotion about our methodology and at the same time warning about the pitfalls and pointing out limitations. In this context a new commandment was formulated: never apply bibliometric analysis as a standalone tool, but always use it in combination with peer review.

1996 – 2000
Impact factor problems; dynamic mapping of research fields; peer review compared with a set of bibliometric indicators; influence of international collaboration on impact; growth, ageing and fractal differentiation of science; patent citation analysis; cross-disciplinary citation flows; better journal impact indicators; measuring scientific excellence; publication delays.

---

2 Maastricht Economic and Social Research Institute on Innovation and Technology, https://www.maastrichtuniversity.nl/research/united-nations-university-maastricht-economic-and-social-research-institute-innovation (July 15, 2020).

3 VSNU: Association of Universities in the Netherlands, https://www.vsnu.nl/en_GB (July 15, 2020).

Due to the steady increase of contract work, it was decided in 2002 to set up *CWTSbv* as a spin-off company of CWTS. This was a very important step in the institutionalization of CWTS. Setting up a spin-off company in a field such as bibliometrics with a well-arranged personnel and financial connection between company and university is no easy task. To find and get support from the right people proved, again, essential. My Leiden colleague astronomy professor, ESO[4] Director-General Harry van der Laan, and LURIS[5] director Ben Hiddinga all played a decisive role in the creation of CWTSbv.

But we still had one final step to go and that would take several years. Meanwhile, the internet had changed scientific communication. More and more data on all types of publications and other research outlets became available in institutional and personal websites. The use of these data, webometrics, and later the use of social media data, altmetrics, provided new opportunities in scientometric research and evaluation studies next to the WoS and, since 2004, Scopus data. As discussed earlier, books and in particular handbooks produced by an institute are, next to publications in international journals, an important sign of institutionalization. In a period of two years two CWTS-based (hand)books appeared (Moed et al., 2004; Moed, 2005).

An important and far-reaching event in the scientometric world was the emergence of *university rankings*, with the Shanghai Ranking[6] as the first in 2003. Shortly after that the Times Higher Education launched its ranking[7] and CWTS introduced the Leiden Ranking.[8]

2001–2005
Effect of language biases in international comparisons; benchmarking scientific excellence; research performance of China; inventor opinions on science dependence of technologies; scientific basis of applied research; Sleeping beauties in science; effects of commercialization of research on public knowledge production; conceptual and methodological problems in university rankings; relation between downloads and citations; inter-field knowledge transfer.

Then, in 2008, the Minister of Education, Culture and Sciences decided to grant CWTS a substantial amount of ear-marked financing in order to improve its innovative power. For the first time CWTS received a permanent direct university funding support of substantial size. This funding made it possible to set up a long-term research program with multiple themes and to initiate PhD work on a larger scale.

4 European Southern Observatory, https://www.eso.org/ (July 15, 2020).
5 LURIS is the Knowledge Exchange Office for Leiden University and Leiden University Medical Center, https://luris.nl/ (July 15, 2020).
6 http://www.shanghairanking.com/index.html (July 15, 2020).
7 http://www.timeshighereducation.co.uk/world-university-rankings (July 15, 2020).
8 http://www.leidenranking.com/ (July 15, 2020).

The then dean of the Faculty of Social and Behavioral Sciences, Theo Toonen, professor of public administration, played a crucial role in the Minister's decision. Also, in 2008, we moved to the renovated Willem Einthoven[9] building with a complete floor of our own. In my opinion, these events in 2008 can be seen as the actual completion of the institutionalization of our research and as the crowning glory of years of efforts to get scientometric research recognized as a scientific field with professional practitioners in an academic context.

2006 – 2010
Measuring research performance in the social science and the humanities; non-source citation analysis; comparison of h-index with crown indicator; effects of Open Access on citation impact; international mobility of Chinese researchers; normalization of co-occurrence data; spatial patterns of scientific collaboration; contextual citation impact; research performance at the individual level, influence of age; VOSviewer mapping tool.

Therefore, I conclude this history of the institutionalization of scientometrics illustrated by the issues that CWTS had to face. Of course, life goes on after 2008, new chairs were created (science policy; science and innovation studies), and this author retired in 2010 as Director of CWTS. Paul Wouters was appointed as Professor of Scientometrics and became the new director. He energetically took over the responsibility for CWTS. New research themes were developed such as responsible evaluation practices, scientific careers, social impact of science, innovation studies, and open science. In particular, the development of new mapping software (van Eck and Waltman, 2010; 2014) and a redefinition of our crown indicator (Waltman et al., 2011a; 2011b) has been of great importance.

2011 – 2015
New crown indicator; inconsistency h-index; publication-level classification; methodology of the Leiden Ranking; percentile-based indicators; clinical research underestimated in citation analysis; community detection in large networks; altmetrics, social media mentions and citation impact; CitNetExplorer; document properties and collaboration.

After the retirement of Paul Wouters in 2018 the triumvirate of Sarah de Rijcke, Ludo Waltman and Ed Noyons is now in charge of CWTS. For the newest developments we refer to the website of CWTS.[10] To conclude, Figure 1 shows a co-word landscape of the recent CWTS activities in scientometric research based on the 2016 – 2019 publi-

9 Our building is named after Willem Einthoven, Professor of Physiology at Leiden, inventor of the electrocardiograph, Nobel Prize 1924.
10 https://www.cwts.nl/ (July 15, 2020).

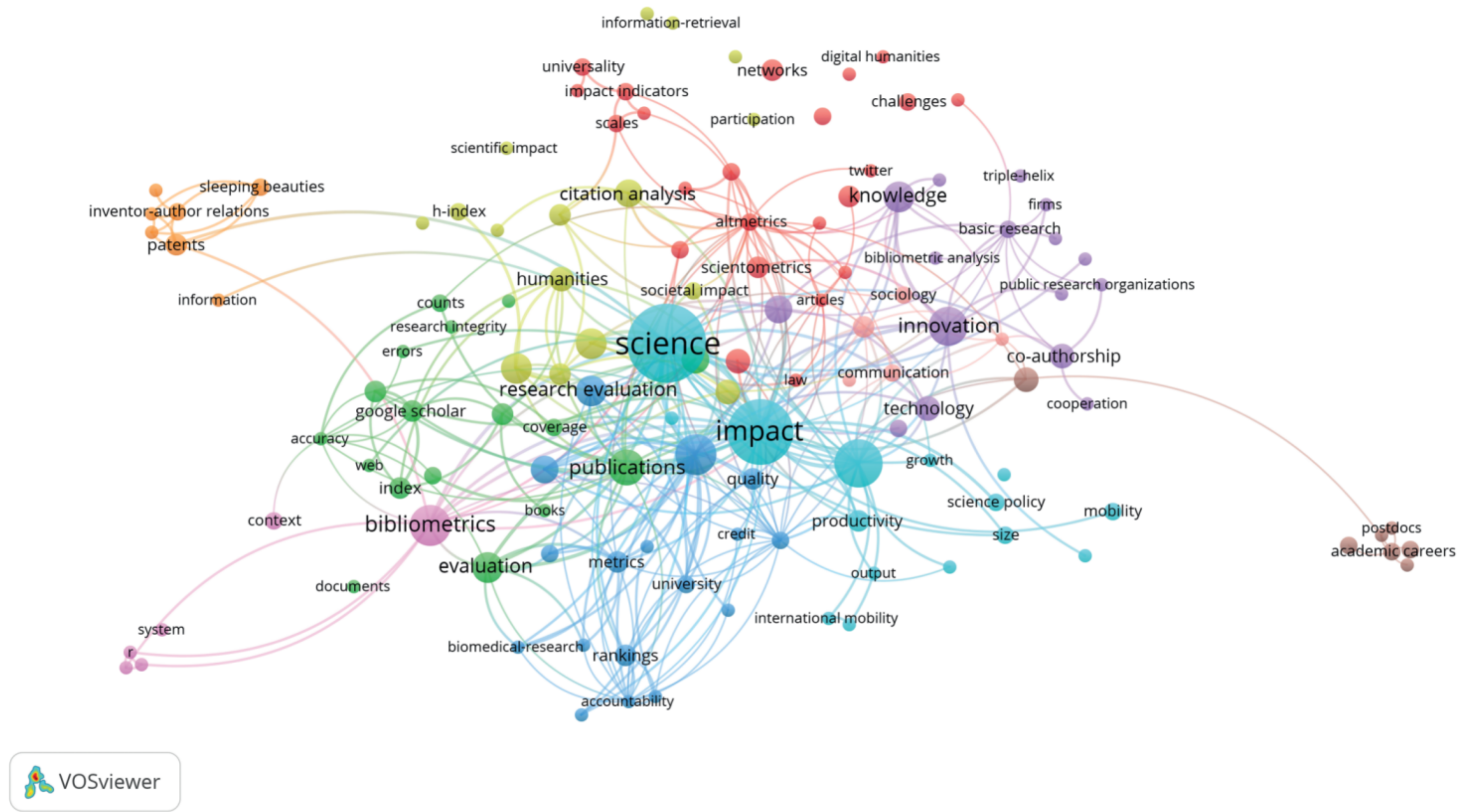

Figure 1: Map of CWTS research in recent years (co-word analysis, occurrence threshold=2, full counting modality).

cations. We clearly see thematic clusters (indicated by different colors) of both the long-standing core issues as well as the new developments.

2016 – 2019
Evaluation practices; citation-based clustering; social media metrics; Google Scholar and evaluation of social science and humanities research; funding acknowledgements; self-citations and impact factor; career attitudes; full versus fractional counting; in-text citations; university-industry R&D linkages.

References

Bakker, CJG 1977, *Elektronenmicroscopie in Nederland*, FOM, Utrecht, FOM-Report 43105.
Callon, M, Bauin, S, Courtial, JP & Turner W 1983, 'From translation to problematic networks: an introduction to co-word analysis', *Social Science Information*, vol. 22, pp. 191 – 235.
Chang, KH 1975, *Evaluation and survey of a subfield of physics: magnetic resonance and relaxation studies in the Netherlands*, FOM, Utrecht, FOM-Report 37175.
Garfield E 1955, 'Citation indexes for science: A new dimension in documentation through association of ideas', *Science*, vol. 122, no. 3159, pp. 108 – 111.
Martin, BR & Irvine J 1983, 'Assessing Basic Research: Some Partial Indicators of Scientific Progress in Radio Astronomy', *Research Policy*, vol. 12, pp. 61 – 90.
Moed, HF, Burger, WJM, Frankfort, JG & Van Raan, AFJ 1983, *On the Measurement of Research Performance: The Use of Bibliometric Indicators*, Science Studies Unit, Leiden.
Moed, HF, Glänzel, W & Schmoch, U (eds.) 2004, *Handbook of Quantitative Science and Technology Research*, Kluwer, Dordrecht.
Moed, HF 2005, *Citation Analysis in Research Evaluation*, Springer, Dordrecht.
Narin, F 1976, *Evaluative Bibliometrics: The Use of Publication and Citation Analysis in the Evaluation of Scientific Activity*, National Science Foundation, Washington D.C.
Over Leven (About Life) (1982). Report of the Verkenningscommissie Biochemie (National Survey Committee on Biochemistry). The Hague: Staatsuitgeverij (in Dutch).
Rigter H (1983). *De prestaties van het Nederlandse gezondheidsonderzoek* (Performance of health research in the Netherlands). RAWB (Council for Science Policy), Series Background Studies, Report 9. The Hague: Staatsuitgeverij (in Dutch).
Rigter H (1986). *Evaluation of the performance of health research in the Netherlands.* Research Policy 15, 33 – 48.
Small, H 1973, 'Co-citation in the scientific literature: a new measure of the relationship between two documents', *Journal of the American Society for Information Science and Technology*, vol. 24, pp. 265 – 269.
van Eck, NJ & Waltman, L 2010, 'Software survey: VOSviewer, a computer program for bibliometric mapping', *Scientometrics*, vol. 84, no. 2, pp. 523 – 538.
van Eck NJ & Waltman, L 2014, 'CitNetExplorer: A new software tool for analyzing and visualizing citation networks', *Journal of Informetrics*, vol. 8, no. 4, pp. 802 – 823.
van Raan, AFJ (ed.) 1988, *Handbook of quantitative studies of science and technology*, North Holland, Amsterdam.
van Raan, AFJ 2004, 'Measuring Science. Capita Selecta of Current Main Issues', in HF Moed, W Glänzel & U Schmoch (eds.), *Handbook of Quantitative Science and Technology Research*, pp. 19 – 50, Kluwer, Dordrecht.

van Raan, AFJ 2019, 'Measuring Science: Basic Principles and Application of Advanced Biliometrics', in W Glänzel, HF Moed, U Schmoch & M Thelwall (eds.), *Handbook of Science and Technology Indicators*, Series: Springer Handbooks, Springer, Heidelberg.
Waltman, L, van Eck, NJ, van Leeuwen, TN, Visser, MS & van Raan, AFJ 2011a, 'Towards a new crown indicator: An empirical analysis', *Scientometrics*, vol. 87, no. 3, pp. 467 – 481.
Waltman, L, van Eck, NJ, van Leeuwen, TN, Visser, MS & van Raan, AFJ 2011b, 'Towards a new crown indicator: Some theoretical considerations', *Journal of Informetrics*, vol. 5, no. 1, pp. 37 – 47.
Wouters, PF 1999, The Citation Culture, PhD thesis, University of Amsterdam.

1.6 International Conferences of Bibliometrics

Grischa Fraumann, Rogério Mugnaini, and Elías Sanz-Casado

Abstract: Conferences are deeply connected to research fields, in this case bibliometrics. As such, they are a venue to present and discuss current and innovative research, and play an important role for the scholarly community. In this article, we provide an overview on the history of conferences in bibliometrics. We conduct an analysis to list the most prominent conferences that were announced in the newsletter by ISSI, the International Society for Scientometrics and Informetrics. Furthermore, we describe how conferences are connected to learned societies and journals. Finally, we provide an outlook on how conferences might change in future.

Keywords: international conferences, bibliometrics, scientometrics, informetrics, altmetrics, history, learned societies, ISSI.

Introduction

Conferences are deeply connected to research fields, in this case bibliometrics. As such, they are a venue to present and discuss current and innovative research, and play an important role for the scholarly community. They also serve as a venue to play a variety of roles, strengthening the so-called invisible colleges (Zuccala, 2006), and are important for gaining scientific reputation (Söderqvist and Silverstein, 1994). Furthermore, conference awards and committee memberships are a marker of prestige among scholars (Jeong, Lee and Kim, 2009). This chapter provides an overview on international conferences in bibliometrics, and what role they play in the history and institutionalization of bibliometrics. Proceedings papers are published in conference proceedings, and such proceedings are also indexed, for example, since 1990 by the *Conference Proceedings Citation Index* (*CPCI)* (Sugimoto and Larivière, 2018) and in *Scopus* (Gingras, 2016). Apart from books and journal articles, proceedings papers have a long tradition in disseminating research (Sugimoto and

Grischa Fraumann, Research Assistant at the TIB Leibniz Information Centre for Science and Technology in the R&D Department, PhD Fellow at the University of Copenhagen in the Department of Communication, Research Affiliate at the "CiMetrias: Research Group on Science and Technology Metrics" at the University of São Paulo (USP), gfr@hum.ku.dk
Rogério Mugnaini, Professor of Library and Information Science at the University of São Paulo (USP), where he leads the research group "CiMetrias: Research Group on Science and Technology Metrics", mugnaini@usp.br
Elías Sanz-Casado, Full Professor in the Department of Library and Information Science at the Carlos III University of Madrid, Director of the research group "Laboratory for Metric Information Studies" (LEMI), leads the "Research Institute for Higher Education and Science" (INAECU) which is made up of members of Carlos III University of Madrid and Autonomous University of Madrid, elias@bib.uc3 m.es

https://doi.org/10.1515/9783110646610-008

Larivière, 2018), and are used as datasets for bibliometric (Glänzel et al., 2006; Lisée, Larivière, and Archambault, 2008) and altmetric studies (Thelwall, 2019). On the one hand, not all conference proceedings are indexed, which makes such citation indexes sometimes incomplete (Sugimoto and Larivière, 2018). On the other hand, some conferences publish their proceedings as journal special issues or books, which can lead to the indexing of proceedings papers. Proceedings papers play a particular role in natural sciences and medicine (Ball, 2017). They represent an important means of publication in computer science (Fathalla et al. 2018), while they might be rather irrelevant in other disciplines, such as sociology (Jeong, Lee and Kim, 2009). As such, they also contribute to faster knowledge sharing than journal articles.

A Historical Sketch on Conferences in Bibliometrics

There are several examples of early conferences in bibliometrics. In 1946, two international conferences on scientific information were the first events, organised by the Royal Society of London (Gingras, 2016). The goal was to develop new forms of indexing scientific literature. This was also related to the exponential amount of scientific literature that led to the foundation of a citation index as part of Garfield's Institute for Scientific Information (ISI) in 1963. The first international conference was held in 1974, titled "Toward a Metric of Science: The Advent of the Science Indicators" (Gingras, 2016). 1987 is considered as the start of a new era in international conferences, since a predecessor of the conferences organised by the International Society for Scientometrics and Informetrics (ISSI) was held for the first time.

This led in 1993 to the foundation of a learned society, namely ISSI (Gingras, 2016). Conferences, journals, and learned societies go in line with a consolidation of an academic discipline, as the foundation of the journal *Scientometrics* in 1978 demonstrates (Gingras, 2016). Most conferences nowadays also experience other forms of scholarly communication, such as the live tweeting about conference presentations and discussions (Holmberg, 2015). Such tweets may also lead to collaborations for researchers that do not physically attend a conference (Holmberg, 2015). Furthermore, conference reports are often also communicated via blogs. Other forms of communication are live streams or archived videos of conference presentations that are available on dedicated online platforms (Plank et al., 2019).

An Overview of International Conferences and its Relation to Learned Societies

ISSI is one of the largest learned societies in bibliometrics, among others. Established in 1993 by a group of researchers during a conference in Berlin, the Netherlands-based association coordinates the ISSI Conference, a members' directory, a blog, and a quarterly newsletter. Conference participants and society members get access to the ISSI conference proceedings, which contain all proceedings papers. ISSI also advocates for several international initiatives, such as the Initiative on Open Citations (I4OC), which has been supported by ISSI in an open letter (Sugimoto, Murray, and Larivière, 2018).

The influence by the society on the research fields can be observed by its significant number of members and conference participants. The conferences used to run independently (ISSI, 2019), and in 2019 were held for the first time together with the European Network of Indicators Designers (ENID). The biennial conference is held in different locations around the world, and the abovementioned first conference in 1987 was called "International Conference on Bibliometrics and Theoretical Aspects of Information Retrieval". As happens often within the scholarly community, anecdotal evidence suggests that the conference was started by a discussion between two researchers, and the question "Shouldn't we start a biennial international conference on informetrics?" (ISSI, 2015). Since 1993 it bears the same name as of today (Hood and Wilson, 2001; ISSI, 2019). The conference is reported to be one of the world's largest and most prestigious conferences (Gorraiz et al., 2014).

Considering the influence of ISSI, the quarterly newsletter is used as a dataset to query past and ongoing international conferences in this research field. A related data collection method has been carried out, for instance, by Söderqvist and Silverstein (1994) for conferences in immunology and Jeong et al. (2009) for conferences in bioinformatics. The *ISSI Quarterly Newsletter* started in 2005, and published 59 issues until September 2019, as of October 28, 2019. All newsletters are publicly available also for non-members. The newsletter in PDF has an ISSN and is curated by 10 members of an editorial board according to guidelines and potential authors need to submit proposals. This is to say the structure is rather similar to a magazine than to a newsletter. Apart from discussions on current research and the introduction of members as well as other news of the society, the newsletter includes conference announcements, call for papers and conference reports. Formats such as workshops, meetings, symposia, forums, summer schools, PhD courses, seminars, and other training courses were excluded from the selection in this article. All quarterly newsletters since the start were downloaded. This approach might have some limitations. For example, certain conferences might be excluded, because they were announced somewhere else. Still, the approach provides a glimpse into the world of conferences on bibliometrics. The conferences were ordered according to the frequency or status, and additional information on their listing in *Conference Proceedings Citation Index (CPCI)* as well as from conference websites was provided if available (see Table 1). The conferences that were discontinued over time are not included in Table 1, except for the UK Social Networks Conference that was merged with two other conferences.

Table 1: International conferences in bibliometrics (*N*=11) ordered by frequency/ status including information on their listing in *CPCI* (source: *ISSI Quarterly Newsletter* March 2005 until September 2019, number 1–59, retrieved from http://issi-society.org/publications/issi-newsletter/ [July 15, 2020]).

Conference name	Frequency/ Status	Listed in CPCI	Description
International Conference "Impact of Science" – Measuring and Demonstrating the Societal Impact of Science	Several times per year	-	"[A] conference […] to discuss measuring and demonstrating the societal impact of science" (https://scienceworks.nl/impact-of-science-2015/ [July 15, 2020])
Triple Helix Conference	Annually	+	"The Triple Helix model presents an opportunity to achieve innovation outcomes for the socio-economic good through collaboration with multi-stakeholders within academia, industry and government spheres." (https://triple-helix.co.za/ [July 15, 2020])
S&T Indicators Conference	Annually	+	Conference on Science and Technology Indicators
InSciT Conference	Annually	-	Conference on Science of Team Science
WissKom conference	Annually	-	Conference of the Central Library at Forschungszentrum Jülich, Germany
CARMA: International Conference on Advanced Research Methods and Analytics	Annually	+	"Research methods in economics and social sciences are evolving with the increasing availability of Internet and Big Data sources of information. As these sources, methods, and applications become more interdisciplinary, the […] International Conference on Advanced Research Methods and Analytics (CARMA) aims to become a forum for researchers and practitioners to exchange ideas and advances on how emerging research methods and sources are applied to different fields of social sciences as well as to discuss current and future challenges." (http://www.carmaconf.org/ [July 15, 2020])
iConference	Annually	+	"[…] insights on critical information issues in contemporary society." (https://ischools.org/iConference [July 15, 2020])

Table 1 *(Continued)*

Conference name	Frequency/ Status	Listed in CPCI	Description
International Conference on Webometrics, Informetrics and Scientometrics (WIS) & COLLNET Meeting	Annually	+	"[...] all aspects of webometrics, informetrics and scientometrics." (http://collnet2019.dlut.edu.cn/meeting/index_en.asp?id=2676 [July 15, 2020])
International Conference on Scientometrics and Informetrics (ISSI)	Biennial	+	"The goal of [the] ISSI [conference] is to bring together scholars and practitioners in the area of informetrics, bibliometrics, scientometrics, webometrics and altmetrics to discuss new research directions, methods and theories, and to highlight the best research in this area." (https://www.issi2019.org/ [July 15, 2020])
Atlanta Conference on Science and Innovation Policy	Biennial	+	"The Atlanta Conference on Science and Innovation Policy provides a showcase for the highest quality scholarship from around the world addressing the challenges and characteristics of science and innovation policy and processes." (http://www.atlconf.org/ [July 15, 2020])
UK Social Networks Conference	Discontinued	-	The UK Social Networks Conference merged with the Applications of Social Network Analysis to form the European Conference on Social Networks (EUSN) (https://www.eusn2019.ethz.ch/ [July 15, 2020])

To the best of the authors' knowledge, there is no comprehensive discipline-specific database available that includes all conferences in bibliometrics, while there are several databases worldwide, for example discipline-specific ones, such as *dblp computer science bibliography* (Ley, 2002). Generally speaking, these databases list proceedings papers that are linked to, for example, authors and conferences. Apart from the conferences mentioned in Table 1, one could add the Altmetrics Conference, an annual conference that provides a venue for research and other initiatives on altmetrics, that is metrics to track research articles online (Priem et al., 2010), but only the related Altmetrics Workshop was mentioned in the *ISSI Quarterly Newsletter*, and workshops were excluded from this selection. Other available databases on a national level might include the Brazilian *Lattes Platform* (Marques 2015) that shows CVs of researchers and their publications as well as the attended national and international conferences (Mugnaini et al., 2019). A similar database might also be provided by institutional, national or international CRIS (Current Research Information Systems) (Sivertsen, 2019), such as *DSpace-CRIS* (Palmer et al., 2014) or *VIVO* (Börner et al., 2012; Conlon et al., 2019). Nevertheless, there are certain ongoing initiatives that might provide in future a more nuanced view on international conferences. One example is the *ConfIDent platform* that is developed as part of a research project funded by the German Research Foundation (DFG). The project objective is to develop a Wiki-based platform that takes into account the needs of scholarly communities, and offers a curated list of conferences. By structuring the conference data according to requirements of interoperability on the technical side and to academic demands on the social side, the system aims to present the possibility for a socio-technical quality assessment of the content (Hagemann-Wilholt, Plank, and Hauschke, 2019; Hagemann-Wilholt, 2019; Sens and Lange, 2019). Generally speaking, proceedings papers may also be linked to ORCID IDs of researchers, that is personal identifiers (Dreyer et al., 2019). There are also other existing online platforms, such as *Open Research* (Fathalla et al., 2019) and *ConfRef*. Additionally, there are further initiatives to develop persistent identifiers (PIDs) for conferences (Crossref, 2019). There are also initiatives underway to develop a semantic representation of scientific events (Fathalla and Lange, 2018) in knowledge graphs, and to make structured queries available to the wider public (Fathalla, Lange, and Auer, 2019a). Such datasets may also be used to rank conferences or explore the impact of these events (Fathalla, Lange and Auer, 2019b; Hansen and Budtz Pedersen, 2018; Hauschke, Cartellieri, and Heller, 2018; Altemeier, 2019).

Conclusions

Conferences are important vehicles to disseminate research and to network with peers. Conferences aligned with journals and learned societies serve an important role in the institutionalization of bibliometrics. Bibliometricians and altmetricians also develop metrics based on data from conferences. There are ongoing initiatives to develop persistent identifiers for conferences. How will the future of conferences look like? Live streaming and other online tools, among others, might have changed the importance of traveling to conferences, and social media makes it possible to join the discussion remotely, but conferences will most probably remain prominent in the academic world, because the social component of meeting other researchers is one of the most important aspects. However, the climate crisis and coronavirus pandemic will most probably require fundamental changes for most academic conferences, and will accelerate digital alternatives and make it less necessary to physically attend these events (Viglione, 2020).

References

Altemeier, F, 'Konferenzmetadaten als Basis für szientometrische Indikatoren', in *VIVO Workshop 2019*, Hannover.

Ball, R 2017, *An Introduction to Bibliometrics: New Development and Trends*, Elsevier Science.

Börner, K, Conlon, M, Corson-Rikert, J & Ding, Y 2012, 'VIVO: A Semantic Approach to Scholarly Networking and Discovery', *Synthesis Lectures on the Semantic Web: Theory and Technology*, vol. 2, no. 1, pp. 1–178, https://doi.org/10.2200/S00428ED1 V01Y201207WBE002.

Conlon, M, Woods, A, Triggs, G, O' Flinn, R, Javed, M, Blake, J & Gross B et al. 2019, 'VIVO: A System for Research Discovery', *Journal of Open Source Software*, vol. 4, no. 39, p. 1182, https://doi.org/10.21105/joss.01182.

Crossref, 'PIDs for Conferences & Projects', https://www.crossref.org/working-groups/conferences-projects/ (July 15, 2020).

Dreyer, B, Hagemann-Wilholt, S, Vierkant, P, Strecker, D, Glagla-Dietz, S, Summann, F, Pampel, H & Burger, M 2019, 'Die Rolle Der ORCID ID in der Wissenschaftskommunikation: Der Beitrag des ORCID-Deutschland-Konsortiums und das ORCID-DE-Projekt', *ABI Technik*, vol. 39, no. 2, pp. 112–21, https://doi.org/10.1515/abitech-2019–2004.

Fathalla, S & Lange, C 2018, 'EVENTSKG: A Knowledge Graph Representation for Top-Prestigious Computer Science Events Metadata', in NT Nguyen et al. (eds.), *Computational Collective* Intelligence, vol. 11055, pp. 53–63, Lecture Notes in Computer Science, Springer International Publishing, Cham.

Fathalla, S, Lange, C & Auer, S 2019a, 'A Human-Friendly Query Generation Frontend for a Scientific Events Knowledge Graph', in A Doucet (ed.), *Digital Libraries for Open Knowledge*, vol. 11799, pp. 200–214, Lecture Notes in Computer Science, Springer International Publishing, Cham.

Fathalla, S, Lange, C & Auer, S 2019b, 'EVENTSKG: A 5-Star Dataset of Top-Ranked Events in Eight Computer Science Communities', in P Hitzler et al. (eds.), *The Semantic Web*, vol. 11503, pp. 427–442, Lecture Notes in Computer Science, Springer International Publishing, Cham.

Fathalla, S, Vahdati, S, Auer, S & Lange, C 2018, 'Metadata Analysis of Scholarly Events of Computer Science, Physics, Engineering, and Mathematics', in E Méndez et al. (eds.), *Digital Libraries for Open Knowledge*, vol. 11057, pp. 116–128, Lecture Notes in Computer Science, Springer, Cham.

Fathalla, S, Vahdati, S, Auer, S & Lange, C 2019, 'The Scientific Events Ontology of the OpenResearch.Org Curation Platform', in C-C Hung & GA Papadopoulos (eds.), *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing – SAC '19*, pp. 2311–13, ACM Press, New York.

Gingras, Y 2016, *Bibliometrics and Research Evaluation: Uses and Abuses*, History and foundations of information science, The MIT Press, Cambridge Massachusetts.

Glänzel, W, Schlemmer, B, Schubert, A & Thijs, B 2006, 'Proceedings Literature as Additional Data Source for Bibliometric Analysis', *Scientometrics*, 68, no. 3, pp. 457–73, https://doi.org/10.1007/s11192–006–0124-y.

Gorraiz, J, Gumpenberger, C, Hörlesberger, M, Moed, H & Schiebel E 2014, 'The 14th International Conference of the International Society for Scientometrics and Informetrics', *Scientometrics*, vol. 101, no. 2, pp. 937–38, https://doi.org/10.1007/s11192-014-1438–9.

Hagemann-Wilholt, S, Plank, M & Hauschke C 2019, 'ConfIDent – for FAIR Conference Metadata: Development of a Sustainable Platform for the Permanent and Reliable Storage and Provision of Conference Metadata', in *21st International Conference on Grey Literature.*

Hagemann-Wilholt, S 2019, *ConfIDent – Eine Verlässsliche Plattform Für Wissenschaftliche Veranstaltungen*, WissKom 2019, Jülich, Germany, 4 Jun 2019–6 Jun 2019, http://juser.fz-juelich.de/record/863207 (July 15, 2020).

Thelwall, M 2019, ‘Mendeley Reader Counts for US Computer Science Conference Papers and Journal articles’, *Quantitative Science Studies*, pp. 1–16, https://doi.org/10.1162/qss_a_00010.
Viglione, G 2020, A year without conferences? How the coronavirus pandemic could change research, Nature 579, 327–328. https://doi.org/10.1038/d41586-020-00786-y.
Zuccala, A 2006, ‘Modeling the invisible college’, *Journal of the American Society for information Science and Technology*, vol. 57, no. 2, pp. 152–168.

2 Theory, Principles and Methods of Bibliometrics

2.1 Peer Review and Bibliometrics

Bernhard Mittermaier

Abstract: Peer review is an established process supporting decisions made on journal publications, grant applications, and tenure, but also helping to assess research groups. For several reasons, peer review is currently being debated, and bibliometrics could serve as its substitute. A large number of studies comparing both approaches has been published, with an overview of their results presented and discussed in this chapter. Although there are good reasons to be hesitant about utilizing bibliometric approaches to assess single persons (e. g. for tenure), the situation is different when assessing research groups. In the STM area, bibliometric indicators could be used as a replacement for peer review.

Keywords: peer review, peer judgement, research assessment, grants, tenure.

Introduction

Peer review is an evaluation of a manuscript or a research proposal, for example. It is carried out by one or more people with expertise similar to that of the producers of the work, i. e. peers (Weller, 2001). The term also includes the retrospective evaluation of the past performance of a scientist or a group of scientists (Gemma, 2017). In a slightly broader sense, the term "peer review" is also used to assess, for example, teaching (Chism and Chism, 2007; Samson and McCrea, 2008; Snavely and Dewald, 2011; van Valey, 2011) or medical staff and practices (Chop and Eberlein-Gonska, 2012; Ertel and Aldridge, 1977; Hadian et al., 2018; Lang, 1999). The latter forms are not discussed here. This chapter opens with an overview of peer review, before presenting a selection of studies comparing peer review and bibliometric indicators. It concludes with a discussion of the findings.

Peer Review

The Development of Peer Review

It is often claimed – probably for the first time by Zuckerman and Merton (1971) – that peer review was invented by the first secretary of the Royal Society, Henry Oldenburg (1610 – 1677). Further occurrences of this narrative can be found in Baldwin (2018), which has led to the impression that peer review is more or less inextricably linked with scholarly publishing. However, Melinda Baldwin showed that it was only

Dr. **Bernhard Mittermaier**, Forschungszentrum Jülich, Central Library, 52425 Jülich, Germany, b.mittermaier@fz-juelich.de

https://doi.org/10.1515/9783110646610-009

over the course of the nineteenth and early twentieth centuries that a number of learned societies adopted the practice of systematically consulting anonymous reviewers about submitted papers. For instance, the Royal Society of Chemistry adopted reviewing systems in the nineteenth century. The American Physical Society implemented this practice in the early twentieth century, though it was not until the 1960s that all submissions to their flagship journal *Physical Review* were peer-reviewed. At this time, the editors of *Nature* still abstained from consulting external reviewers if papers were submitted or recommended by scientists whom they trusted. It was not until 1973 that external peer review became mandatory for manuscripts submitted to *Nature* (Baldwin, 2015). As late as 1989, an editorial in *The Lancet* revealed a huge inner distance to peer review:

> In the United States far too much is being demanded of peer review. Careers and the viability of whole departments now depend on publication in peer-reviewed journals. In the public domain the process is sometimes seen as a guarantee of truth, which is silly; (...) Journals do things differently, and long live those differences, but there was consensus that turning away papers within the editorial board or 'in house' without an outside opinion by no means disqualified a journal from calling itself peer reviewed and that reviewers are advisers (always The Lancet's preferred term) not decision makers. (Anonymous, 1989)

The Peer-review Process

Probably the most important type of peer review occurs when a manuscript is submitted to a journal (Paltridge, 2017). Journals that have implemented a peer review process are referred to as "peer-reviewed journals" or "refereed journals"; sometimes only publications in such journals are considered "real" scientific output. The first step in the process after submission is assessment by the editor. At this stage, some submissions are rejected ("desk rejection") because they are of low quality or because they are beyond the journal's scope. Once they pass this initial screening, articles are sent to external reviewers. In some cases, authors are invited to suggest reviewers. The number of reviewers ranges from one to three or even more at the discretion of the editor or depending on the regulations of the journal. The standard procedure is "blind peer review", where the author receives the reviewers' comments, but the reviewers themselves remain anonymous. In "double-blind peer review", the author is also anonymized. However, this is not always an easy task, for instance if information about the author's affiliation is necessary to understand the article. The rationale here is to enable reviewers to express their opinions openly and, in the case of double-blind peer review, to avoid any bias with regard to the authors' gender, age, reputation, etc. "Open peer review" is the opposite approach, applied for example by journals published by Copernicus. Here, manuscripts are published in their initial form and then undergo a public review (everybody can comment) in addition to invited and more in-depth reviews. All comments can be read free of charge.

In most cases, reviewers are asked to express their opinion in a structured manner. The intention is to help the editor in judging the cmments and the authors in implementing them. Most often, reviewers have to opt for one final recommendation: for example, “accept as is”; “accept with minor revisions”; “accept with major revisions”; or “reject”. If – in the case of two reviewers – there is a substantial disagreement between the recommendations, editors either make the final decision themselves or invite a third reviewer.

If a revision is deemed necessary, authors are advised to follow the reviewers’ suggestions, though it is possible in principle for the authors to discuss changes they may consider unreasonable with the editor. It is common practice for authors to provide a detailed response to each reviewer, explaining how they have followed the reviewer‘s advice or, if they have not, their reasons for not implementing certain suggestions. After revision and resubmission, the editor again decides whether the article will be published or not. Overall, rejection rates differ hugely between journals, e.g. between 2% and 68% in the atmospheric sciences (Schultz, 2010).

The peer review process for books is rather heterogeneous (Goldfinch and Yamamoto, 2012). External peer review by one reviewer (or even more than one) is often substituted by an editorial review, particularly in the case of edited books, where a number of different authors write the chapters. In the case of monographs and particularly in the case of textbooks, the publisher provides the editing through his staff.[1]

Grant proposals can be reviewed in exactly the same way as journal articles, with the exception that here reviewers are not asked to improve the text of the proposal but rather to suggest a different experimental setting etc. Often, grant proposals are judged by a review panel in quite a different setting: the reviewers do not act independently of each other but discuss the proposal together after questioning the applicants. This process often comprises a single session in which a number of proposals for the same tender are reviewed. As a grant application not only includes the research proposal itself but also curriculum vitae of the participating scientists along with their publication records and letters of recommendations, the review also takes the past performance of the applicants into account.

An even greater shift towards past performance assessment occurs in the evaluation of research groups or institutions. For instance, in Germany, universities participating in the German federal and state governments’ Excellence Initiative undergo a rigorous assessment carried out by scientists from abroad. The same applies to the extramural research institutions of the Leibniz Association and the Helmholtz Association. They are evaluated every seven years to decide whether they can continue as a member of the Leibniz Association or, in the case of Helmholtz, to what extent the

1 In contrast to this, journal editors are usually not staff members of the publisher, but scientists who pursue this work in addition to their research and teaching duties. They get no or only minimal compensation for their editorial duties (de Knecht, 2019).

image
not
available

image
not
available

image
not
available

two applicants in a given recruitment process were compared using a combination of two bibliometric indicators (an h-index variant and the percentage of highly cited documents), the outcome of the ranking of those two applicants by peers could be predicted in 75% of cases. Jensen, Rouquier, and Croissant (2009) explored the correlation between bibliometric indicators (h index, h index divided by "scientific age", number of citations, number of publications, and average number of citations per publication) and the results of a peer review process concerning the promotion of about 600 researchers at France's Centre Nationale de la Recherche Scientifique (CNRS). The authors found that

> no single indicator is the best predictor for all disciplines. Overall, however, the Hirsch index h provides the least bad correlations, followed by the number of papers published. It is important to realize however, that even h is able to recover only half of the actual promotions. The number of citations or the mean number of citations per paper are definitely not good predictors of promotion. (Jensen, Rouquier, and Croissant, 2009).

Evaluation of Research Groups

There are several studies comparing peer review of research groups with bibliometric indicators. In general, these studies found a (sometimes weak) positive correlation. Wouters et al. (2015) explain "the imperfect correlations between bibliometric indicators and peer review (partly) by variation in qualitative peer-based judgements". For example, Aksnes and Taxt (2004) compared the peer ratings of 34 research groups at the University of Bergen (Norway) with a set of five bibliometric indicators. The highest Pearson's correlation was observed between peer ratings and an indicator called "relative publication strategy". It compares the average citation rate of the journals in which the group's articles were published with the average citation rates of the subfields covered by each journal. Meho and Sonnenwald (2000) analyzed the relationship between citation ranking and peer evaluation in assessing senior Kurdologists' research performance. Normalized citation ranking and citation content analysis were highly correlated with peer ranking, both for high-ranked and low-ranked senior scholars. Anthony van Raan et al. performed a number of investigations in this area: Nederhof and van Raan (1993) analyzed the relationship between bibliometric indicators and peer review for six research groups in economics. Peer review and bibliometric findings were generally in agreement. Rinia et al. (1998) showed the correlation between different bibliometric indicators and the outcomes of peer review made by expert panels of physicists in the Netherlands. In the field of physics, they assessed a set of 56 research programmes with approximately 5,000 publications and 50,000 citations. They found the strongest correlation to be between bibliometric indicators and the judgement of the researchers and the research team. Later, van Raan investigated the correlation between standard bibliometric indicators and peer judgement for 147 chemistry research groups in the Netherlands (Van Raan, 2006). He found that both h index and CPP/FCSm discriminate very well between the

image
not
available

average scores, which, according to Traag and Waltman (2019), are more relevant for the REF.

Conclusion

Overall, most of the comparative studies found a moderately positive correspondence between peer review and bibliometric indicators, but the correlations identified have been far from perfect and have varied among the studies. Inter alia, the correlations depend on the scientific field, the bibliometric indicators, and the subject of the review:

- The results of studies focusing on grant decisions are mixed. While a number of studies revealed a positive correlation between grant peer review and citation impact, other studies showed no or only a low correlation between the success in grant applications and subsequent citation impact.
- The results of investigations on tenure decisions are not convincing either. At best, bibliometric indicators can predict the correct ranking of any two applicants in 75 % of cases, which is only halfway between the actual result (100 %) and a random decision (50 %). Reviews of research group assessments generally revealed better correspondence between peer review and bibliometrics, often depending on the scientific field and the indicator in question.
- Investigations on national research assessments revealed results similar to the studies on research group assessments. As a rule of thumb, correlations in the area of science, technology, and medicine (STM) are better than in the social sciences and humanities, and correlations are better for field-normalized indicators than for basic indicators like the citation count.

Therefore, there is generally little empirical support for the hypothesis that bibliometrics reflects the same aspects of impact or research quality as peer review. However, the extent to which the correlation between the two approaches is considered sufficient depends on the nature and the goals of the evaluation. The statement by Abramo and D'Angelo (2011) regarding national research assessments could hold true for the evaluation of research groups as well: "Accepting that there is no one infallible evaluation method, the position of the authors is that for the natural and formal sciences, the bibliometric methodology is by far preferable to informed peer review." This may first appear as a daring thesis, but it can be justified for the following reasons:

- Peer review is far from perfect, as was shown, for example, in the section "Peer Review Versus Journal Impact Factor". Therefore, a deviation in the results of a bibliometric approach from a peer decision does not necessarily indicate that the bibliometric approach led to a "wrong" result.

Walsh, E, Rooney, M, Appleby, L & Wilkinson G 2000, 'Open peer review: A randomised controlled trial', *British Journal of Psychiatry*, vol. 176, no. 1, pp. 47–51, https://doi.org/10.1192/bjp.176.1.47.

Warner, J 2000, 'A Critical Review of the Application of Citation Studies to the Research Assessment Exercises', *Journal of Information Science*, vol. 26, pp. 453–460, https://doi.org/10.1177/016555150002600607.

Weller, AC 2001, *Editorial peer review: its strengths and weaknesses.ASIST monograph series*, Information Today, Medford, N.J.

Wilson, JD 1978, 'Peer review and publication. Presidential address before the 70th annual meeting of the American Society for Clinical Investigation, San Francisco, California, 30 April 1978', *The Journal of Clinical Investigation*, vol. 61, no. 6, pp. 1697–1701, https://doi.org/10.1172/JCI109091.

Wouters, P, Thelwall, M, Kousha, K, Waltman, L, de Rijcke, S, Rushforth, A & Franssen T 2015, *The Metric Tide: Literature Review (Supplementary Report I to the Independent Review of the Role of Metrics in Research Assessment and Management)* (HEFCE).

Zuckerman, H & Merton RK 1971, 'Patterns of evaluation in science: Institutionalization, structure and functions of the referee system', *Minerva*, vol. 9, no. 1, pp. 66–100, https://doi.org/10.1007/bf01553188.

image
not
available

image not available

Index