Springer Handbook of Science and Technology Indicators

Springer Handbooks provide a concise compilation of approved key information on methods of research, general principles, and functional relationships in physical and applied sciences. The world's leading experts in the fields of physics and engineering will be assigned by one or several renowned editors to write the chapters comprising each volume. The content is selected by these experts from Springer sources (books, journals, online content) and other systematic and approved recent publications of scientific and technical information.

The volumes are designed to be useful as readable desk book to give a fast and comprehensive overview and easy retrieval of essential reliable key information, including tables, graphs, and bibliographies. References to extensive sources are provided.

Springer Handbook of Science and Technology Indicators

Wolfgang Glänzel, Henk F. Moed,
Ulrich Schmoch, Mike Thelwall (Eds.)

With 279 Figures and 223 Tables

Springer

Editors
Wolfgang Glänzel
ECOOM and Faculty of Economics and Business
KU Leuven
Leuven, Belgium

Henk F. Moed
Amsterdam, The Netherlands

Ulrich Schmoch
Competence Center Policy – Industry – Innovation
Fraunhofer Institute for Systems and Innovation Research ISI
Karlsruhe, Germany

Mike Thelwall
Faculty of Science and Engineering
University of Wolverhampton
Wolverhampton, UK

ISBN 978-3-030-02510-6 e-ISBN 978-3-030-02511-3
https://doi.org/10.1007/978-3-030-02511-3

This Springer imprint is published by the registered company Springer Nature Switzerland AG, part of Springer Nature.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The *Springer Handbook of Science and Technology Indicators* offers a collection of state-of-the-art contributions on quantitative science and technology research. Organized in six parts, the individual chapters focus on various aspects of the development and application of indicators derived from data on scholarly publications, patents, and electronic communication. The 44 chapters are written by leading specialists in the topics selected for this Springer Handbook. These chapters deal with theoretical and methodological issues, illustrate applications, highlight their policy context and relevance, and point to future research directions. In particular, the authors present a survey of the research topics they address, and show their most recent achievements and contribution to the advancement of quantitative studies of science and technology.

The chapters are arranged into six parts:

Part A: Analysis of data sources and network analysis
Part B: Advancement of methodology for research assessment
Part C: Science systems and research policy
Part D: New indicators for research assessment
Part E: Advancement of methodology for patent analysis
Part F: Patent system, patents, and economics.

The Editors' Introduction provides a further specification of the handbook's scope and of the main topics addressed in its chapters. This Springer Handbook aims at four distinct groups of readers: practitioners in the field of science and technology studies; research students in this field; information scientistsa and practitioners in informatics; scientists, scholars, and technicians who are interested in a systematic, thorough analysis of their activities; policy-makers and administrators who wish to be informed about the potentialities and limitations of the various approaches and about their results.

The current handbook can be considered a successor of the *Handbook of Quantitative Science and Technology Studies* edited by Anthony van Raan and published in 1988 and the *Handbook of Quantitative Science and Technology Research. The Use of Publication and Patent Statistics in Studies of S&T Systems* edited by Henk F. Moed, Wolfgang Glänzel, and Ulrich Schmoch in 2004.

We are grateful to all contributors for their enormous efforts to share their longstanding experience as experts in their research topics and to provide us with excellent chapters for this handbook.

Wolfgang Glänzel
Henk F. Moed
Ulrich Schmoch
Mike Thelwall

Editors' Introduction

The *Springer Handbook of Science and Technology Indicators* continues the tradition and scope set by two predecessor reference works: The *Handbook of Quantitative Studies of Science and Technology*, edited by Anthony F.J. van Raan and published in 1988, and a little more than 15 years later, the *Handbook of Quantitative Science and Technology Research* (editors: Henk Moed, Wolfgang Glänzel, Ulrich Schmoch). Similarly to the previous volumes, this handbook deals with quantitative studies of the science and technology (S&T) system which is conceived as a part of the various national or regional innovation systems.

The current work provides the state of the art of the development and application of methods and models that have been developed to understand and study processes and networks of scientific communication, the indicators for the S&T system that are derived from the documented output of research and patenting activities. Besides reporting and reviewing methodologies and application fields, there is a strong focus on tracing the developments that took place in the field of quantitative S&T studies since the first and second volumes appeared. The three decades since the publication of van Raan's Handbook were characterized by increasing internationalization, the duality of global collaboration and competition in science and technology, challenges to S&T studies that have also created opportunities and proved to be one of the main driving forces for the advancement of our field, and the gradual supplementation and replacement of traditional publishing by electronic communication media and publication channels. Several of these and related issues had already found their way into the second handbook, including the science-technology interface, research collaboration and network analysis, emerging economies and the changing global research landscape, the internationalization of knowledge generation, and data and text mining for S&T studies and webometrics.

The new millennium has sped up development and brought new challenges but also solutions. Increased computing capacity, practically infinite storage capacities, and the development of new algorithms have helped researchers to cope with the challenges of big data that emerged during the last decade. Both the quantity and quality of data now allow the analysis and linkage of huge document corpora, large-scale text mining, and study of the evolution of huge document and actor networks. Open access and open science have improved accessibility to research results and broadened the usage of published information, while scientific blogging provided a platform to communicate science to major stakeholders and the public. The extension of communication and publication channels with new actor and user groups introduced the necessity and possibility of building new measures of usage exchange and networks, which makes it possible to examine new kinds of impact, and to study societal impact beyond the traditional scholarly domain. With the growth of user groups and communities and the wide scope of data sources for information transfer and usage the demand for connectability and interoperability with the necessity of harmonization, standardization, and integration of data emerged (Daraio and Glänzel, 2016).

Software development in recent decades and publicly available data sources, such as Google Scholar and academic licenses for commercial abstract and citation databases with bibliometric features, have opened up bibliometrics to a broader user group among scientists, librarians, and bibliometric semiprofessionals. This has fostered uninformed use of bibliometrics and technometrics, most notably in an evaluative context. This development underlines the necessity of providing an up-to-date handbook on quantitative S&T research to scientists and practitioners, not only reporting the state of the art in the discipline but also giving guidance to practitioners and potential users of S&T indicators.

The contributions to the *Springer Handbook of Science and Technology Indicators* reflect a wide variety of attributes of the contemporary S&T system. Most central concepts have not changed since the previous edition as these include scientific or technological performance, and the productivity or efficiency of the S&T system and its constituent parts. The question of how performance or productivity could be measured also remains a crucial issue but contexts and applications have created new challenges, data provide more and better information, the general trend towards the meso and increasingly micro level has continued, requiring more accuracy and finer granularity. The question of how the various parts in an S&T system react with one another and how this interaction affects the overall performance is still of foremost importance. The need for measuring the impact of research beyond scholarly communication, including policy impact and impact on society, called *broader impact*, has become one of the driving forces for the development of our field. This also includes the necessity of exploiting and mining unstructured and nonstandardized sources for relevant information, a new challenge of big-data processing that has broadened our field further towards computer and information science with the development of new algorithms and new retrieval techniques. These sources and techniques are not replacing the traditional ones based on well-standardized bibliographic and patent databases but they essentially extend traditional quantitative science and technology studies by giving them a new perspective and dimension. To stay abreast of these changes, this handbook includes a new part on the development and application of new indicators (Part D: New Indicators for Research Assessment).

Many new web indicators have been developed since the last handbook was published. In addition to a small set of webometrics, the Altmetrics initiative, started by Jason Priem in 2010, led to the creation of a wide range of new indicators derived from social web sources. It also led to the creation of organizations devoted to collecting and selling altmetric values, including ImpactStory, Altmetric.com, and Plum Analytics. Altmetrics are now widely deployed by publishers in their digital libraries alongside citation counts and are being considered for (limited) research evaluation contexts. Their promise is that they may reflect nonscholarly types of impacts that are ignored by citations from other journal articles, and/or that they may appear more quickly than citations, allowing earlier impact evidence. Most alternative indicators are also easily manipulated and subject to irrelevant uses, however. This makes them difficult to use in most research evaluation contexts. Nevertheless, they can be valuable for self-evaluation and ongoing monitoring purposes (e. g., by funding organizations), if used carefully. They may also be useful on an ad-hoc basis to support narrative claims for nonscholarly impacts. For example, there are instances of this type of use in the case study parts of the 2016 UK Research Excellence Framework evaluations.

In terms of individual new indicators, counts of readers in the social reference manager Mendeley are worth a special mention. These are like citation counts in that they primarily reflect scholarly impact but appear a year earlier, allowing timelier evaluations. In contrast, Altmetric.com scans many policy documents for citations to academic research. Policy mentions give clear evidence of important nonscholarly impacts for academic research. There are also alternative indicators that reflect arts and humanities impacts (e. g., Google Books citations, online reviews), educational impacts (syllabus mentions), and commercial impacts (e. g., Google Patents citations). One of the best-known indicators, tweet counts, appear to primarily reflect publicity, however, and there is no alternative indicator yet that gives good evidence of public interest in research. For this, media mentions are probably the best proxy available. Indicators like download counts are particularly useful for investigating the scholarly communication process by giving insights into patterns of use for documents, including by scholars and students.

In the following, we will give an overview of the chapters that are included in this handbook.

Part A: Analysis of Data Sources and Network Analysis

The chapter by *Vincent Larivière* and *Cassidy R. Sugimoto* deals with the journal impact factor, probably the most widely used bibliometric construct. They argue that this indicator is, by far, the most discussed bibliometric indicator, and has been widely dissected and debated by scholars of every disciplinary orientation. Their chapter presents a brief history of the indicator and highlights a series of well-known limitations, and draws on the existing literature as well as on original research. They highlight the adverse effects of the use of this indicator on authors and publishers, and describe alternative journal-based indicators. Their chapter ends with a call for responsible application of journal indicators, and a commentary on future developments in journal indicators.

Subject delineation has become one of the prominent tasks in bibliometric domain studies. *Michel Zitt*, *Alain Lelu*, *Martine Cadot*, and *Guillaume Cabanac* describe this complex task using three models as a question of disciplines versus invisible colleges. The models, which can be favorably combined with each other, are existing classification schemes, information retrieval, and clustering exercises. The authors discuss the opportunities and limitations of the bibliometric techniques underlying information retrieval, data analysis, and network analysis. They show how multiple network approaches allow the comparison and combination of bibliometric networks. The authors focus on textual and citation networks, but outline possibilities and methods for hybridization. The chapter concludes with a discussion of typical subject delineation schemes and protocols.

In their chapter, *Ronald Rousseau*, *Lin Zhang*, and *Xiaojun Hu* provide a systematic review of interdisciplinarity in scientific research. First, they examine the phenomenon of interdisciplinary research (IDR) from a conceptual perspective and discuss its characteristics and driving forces. The second part is devoted to its quantification and measurement from the information science perspective. The authors proceed from the assumption that IDR is mirrored in the published research documents by the integration of knowledge from different subjects. The authors review existing approaches to measure knowledge integration and show their limitations. Proceeding from the notion by Stirling, Rafols, and Meyer proposing three main components of interdisciplinarity (diversity, variety, and evenness), they propose a true diversity measure. An example from the field of synthetic biology provides an illustration and the chapter concludes with suggestions for future research.

Emilio Delgado López-Cózar, *Enrique Orduna-Malea*, and *Alberto Martín-Martín* argue that the launch of Google Scholar (GS) marked the beginning of a revolution in the scientific information market, because of its automatic indexing of information directly obtained from the web, its ease of use, and its wide coverage. Their chapter lays the foundations for the use of GS as a supplementary source for scientific evaluation, by giving an overview of how GS works, and providing statistics about its size, coverage, and growth rate. In addition, the authors conduct a systematic analysis of the main limitations of GS as an evaluation tool, and compare GS with traditional citation indexes. They conclude that GS presents a broader view of the academic world than the traditional indexes as it includes many previously invisible sources.

The next chapter is devoted to the analysis of current publication trends in gold Open Access (OA). In the first part, *Daniel Torres-Salinas, Nicolas Robinson-Garcia*, and *Henk F. Moed* give a comprehensive review of the current literature on Open Access, specifically in relation to its "citation advantage." This chapter has three dimensions: countries, subject fields, and journals. In the light of this, the authors investigate national gold OA publication patterns, OA journal characteristics and citation differences between gold OA and nongold OA publications, and factors that may affect differences in citation impact between OA and conventional, toll-access journals. The authors also discuss scientists' OA publication strategies and patterns as well as the role of predatory OA journals.

Forecasting future developments in science, technology, and innovation (STI) is the subject of the chapter by *Katy Börner* and *Staša Milojević*. Such forecasts are based on advanced mathematical-statistical and computational models of the STI system, and are facilitated by advances in computational power and in the availability of numerous "big" datasets containing not only bibliometric, but also funding, stock market, social media, and other types of data. Advanced models can be used to simulate and understand the structure and dynamics of the STI system, and enhance human decision-making.

Science mapping in the form of studies of structural aspects of document and actor networks plays an important role in quantitative science studies. The following three chapters tackle this important topic within contemporary scientometric research. The first gives an overview of the advanced bibliometric tool for structural analysis and visualization. The second chapter is devoted to the challenges of the analysis of large-scale bibliometric networks and the third deals above all with fundamental methodological questions of science mapping and topic identification.

Science mapping involves the development and application of computational techniques for the visualization, analysis, and modeling of scientific and technological activities. It is an advanced bibliometric tool to analyze and mine scientific output. *Jose A. Moral-Munoz*, *Antonio G. López-Herrera*, *Enrique Herrera-Viedma*, and *Manuel J. Cobo* review six freely available, comprehensive science mapping tools: Bibexcel, CiteSpace II, CitNetExplorer, SciMAT, Sci2 Tool, and VOSviewer. The authors discuss the strengths and limitations of these tools related to data processing, analysis options, and visualization. They argue that each tool has different properties, and the choice of one over another depends on the type of actors to be analyzed and the type of output expected.

Kevin Boyack and *Richard Klavans* point to the new challenges that have emerged since the last *Handbook of Quantitative Science and Technology Research* was published. The recent science mapping evolution has been facilitated by the availability of full text databases, increased computing capacity, and the development of new algorithms. This has allowed mapping technology to transition from the analysis of small networks to large-scale exercises. The focus is on the analysis of large-scale, global bibliometric networks. The authors give a state-of-the-art report and discuss the commonly used data sources and methods from a historical perspective, continuing to the most recent developments. Their own large-scale topic-level model is used to illustrate the analysis of large-scale bibliometric networks and potential applications.

In his chapter, *Bart Thijs* identifies three drivers of scientometric mapping of science: information-technological innovation; improved community detection; and methodological advancements in the field of scientometrics itself. The author shows that scientometric methodologies using citation-link and lexical approaches lagged the development of the first two drivers. He discusses methodological issues related to community detection. The different approaches to the creation of global maps and the possibility of achieving comparable results at higher levels of granularity are contrasted with the fine-grained solutions possible from local mapping.

Part B: Advancement of Methodology for Research Assessment

Anthony F.J. van Raan gives a comprehensive overview of the methodology and application of advanced bibliometric indicators and introduces bibliometrics as a powerful instrument for the study of science. His historical review starts from the beginning of professional bibliometrics and covers the role of citation indexing in the emergence of the discipline of scientometrics. The review discusses how citation indexing revolutionized quantitative science studies and continues until the stage of contemporary bibliometrics in the internet age. This introduction is followed by a description of advanced state-of-the-art bibliometrics with its rationale and practical needs. The author proceeds from the two main pillars, citation analysis and mapping of science, which can be reduced to a single principle. The author deduces a set of main indicators to be used for research performance assessment with regard to the developments at the Leiden Institute. The conceptual-methodological part is followed by applications of indicators in an evaluative context with various real-life examples. In this context, the author discusses also problematic and controversial issues, such as the use of journal impact factors, the h-index, publication assignment, subject delineation, and university rankings. The last part of the chapter deals with the above-mentioned second pillar of bibliometrics, the mapping of science. Hybrid techniques, the combination of citation analysis and science mapping, and new fields of application are described and discussed.

Ludo Waltman and *Nees Jan van Eck* present a comprehensive overview of a class of bibliometric indicators that are among the most important in bibliometrics, namely field-normalized indicators. The term field indicates a branch of knowledge, such as a research discipline, specialty, or topic. Field-normalized indicators make corrections for differences among fields, so that groups of researchers from different fields can be compared with one another. The authors give an overview of the various field-normalization approaches. Most importantly, they also illustrate how indicators themselves can be evaluated, and how the choice of an approach may affect the outcomes of a bibliometric analysis.

The h-index and its derivatives have become perhaps the most popular and most commonly used bibliometric indicators besides the journal impact factor. Research and applications of Hirsch-type indexes have consequently yielded a large body of literature within our field over the last decade. *András Schubert* and *Gábor Schubert* provide a guided bibliometric tour through more than 3000 papers on this topic. Special attention is paid to the theoretical, mathematical and axiomatic background and various applications as well as the possibility of applying the h-index as a network indicator.

The method of Characteristic Scores and Scales (CSS) was originally proposed in the second half of the 1980s, when their large-scale calculations were still a computational challenge. Because of increased data availability and computational capacity, the method has now become practical. *Wolfgang Glänzel, Bart Thijs*, and *Koenraad Debackere* provide an overview of the various fields of application of this method, which aims to replace the traditional linear approach to citation impact evaluation by a distributional one with a focus on the high end of performance. A discussion of the mathematical background and statistical properties is followed by the implementation of the method in assessment exercises at different levels of aggregation as well as in various disciplinary and multidisciplinary contexts.

The development and application of bibliometric indicators of research performance at the level of individual authors is one of the most debated and complex issues in quantitative science and technology studies. *Lorna Wildgaard* presents a critical overview of the development of this type of indicator. She discusses characteristics and mathematical properties of 68 author-level indicators, and highlights their potential and limitations. The major theme of her contribution is setting the argument for the need to monitor and evaluate current indicator production.

Policy implementation of relevant science, technology, and innovation indicators requires appropriate data management methods, and data integration has become a central issue in this regard. Two main approaches to data integration are in use: procedural and declarative. *Maurizio Lenzerini* and *Cinzia Daraio* follow the latter approach by focusing on the ontology-based data integration (OBDI) paradigm. They discuss the five main principles of this paradigm and the challenges of data integration. Finally, *Sapientia* (the ontology of multidimensional research assessment and its OBDI system) developed at Sapienza University of Rome is provided as an example of an open and collaborative platform for research assessment.

Synergy in innovation systems is studied by *Loet Leydesdorff*, *Inga Ivanova*, and *Martin Meyer* within the framework of the Triple Helix model of university–industry–government relations. This is used as a metaphor in modeling the knowledge-based economy and innovation. Synergy is introduced and analyzed here in the context of the generation of redundancy, the measures of which are derived from an information-theoretic model. Using examples from several countries, it is shown how the Triple-Helix synergy indicator can be applied to analyze regions or sectors in which uncertainty has been significantly reduced and which contribute most to the generation of redundancy. The model and its indicators thus allow the quantification and measurement of the quality of innovation systems at different geographical scales and in terms of sectors.

Part C: Science Systems and Research Policy

The interrelationship between scientometrics and research policy is studied and discussed by *Koenraad Debackere*, *Wolfgang Glänzel*, and *Bart Thijs*. Scientometrics is shown to be a discipline that emerged from the library and information needs of scientific communities and grown into a powerful instrument providing advanced tools and indicators for policy-relevant research assessment. This development is depicted as a symbiosis between scientometrics and science policy. The authors use the example of the *Flemish Expertise Center for R&D Monitoring* (ECOOM) to illustrate this coevolution, pointing to its opportunities, challenges, and limitations.

Research assessment exercises monitoring and evaluating national or regional research performance have a high priority in research management and national research policies. In their chapter, *Sybille Hinze*, *Linda Butler*, *Paul Donner*, and *Ian McAllister* use bibliometric tools to analyze and compare the effects and efficiency of the research assessment regimes of three selected countries (UK, Australia, and Germany). Although the assessment systems of the three countries differ considerably, large differences could not be found regarding their effects and efficiency. They conclude that the systems make less difference than the implementation of an assessment exercise. They further conclude that to understand the mechanisms behind changing performance, indicators are not enough and need to be supplemented by contextual information at various levels of aggregation.

The globalization of research and the use of bibliometric indicators to study this process are the subject of a chapter by *Jacqueline Leta*, *Raymundo das Neves Machado*, and *Roberto Mario Lovón Canchumani*. Given the growing importance of the BRICS countries Brazil, Russia, India, China, and South Africa in the global economy and the science system, the authors focus on scientific collaboration among these countries. They also illustrate how bibliometric techniques can be used to examine traces of the effects of the foundation of the BRICS group upon the international collaboration among its members. A series of techniques was used, including diachronic analysis, Bradford's law, and journal cocitation analysis.

As China publishes over 5000 scientific-scholarly journals, it has developed extensive expertise in journal publishing and journal evaluation. *Zheng Ma* reviews the development of the Chinese journal system in scientific, technical, and medical (STM) fields. The author describes the characteristics of evaluation systems of national journals as compared to those related to international periodicals, in terms of their respective evaluation purposes, evaluation methods, key features, and evaluation criteria. Two cases are presented of China's research work on the evaluation of STM journals, namely the development of the so-called boom index and of comprehensive performance scores for Chinese STM journals. The author also presents analyses of the English-language STM journals in China, and introduces an atomic structure model for evaluating English-language scientific journals published in non-English countries.

Gali Halevi focuses on a crucial issue in science policy, namely the gaps between men and women in the domain of science and scholarship. She provides a thorough review of the various approaches combining bibliometric and other types of research information to the identification of gender among authors of scientific-scholarly literature, and to the measurement of gender disparities. She discusses a series of studies explaining barriers to female participation, and argues that for a comprehensive picture of the underrepresentation of women, bibliometric studies have become an essential tool for tracking not only research participation itself, but also its impact on scientific discovery.

Two chapters are devoted to the measurement of research impact beyond scholarly communication. The first chapter shows how the medical literature is used by clinicians and by the public, while the second reviews and discusses societal impact indicators in recent literature.

The study presented by *Elena Pallari* and *Grant Lewison* analyses how biomedical research could influence its two main goals in improving healthcare: better patient treatment and prevention of illness. They examine two approaches: the research base underlying clinical practice guidelines (CPGs) linked to patient treatment, and stories in the mass media as an expression of healthcare policy. The authors collected CPGs and newsletters from 21 and 22 European countries, respectively, and used Web of Science (WoS) journal articles as their evidence base. The medical research stories from newspapers were linked to research by the WoS papers they cited. The authors found a discrepancy between the papers cited by CPGs and in newspaper stories, on one hand, and those that are frequently cited in scholarly literature, on the other hand. They found that even relatively neglected subject areas could be an important source for medical practice and the general public.

Lutz Bornmann and *Robin Haunschild* give an overview of the literature on societal impact measurement. They first delineate the concept of societal impact, describe the reasons for its emergence, and point to the problems in measuring this kind of impact. Using examples of major projects, they illustrate how frameworks for the measurement of societal impact can be integrated into evaluative contexts. In the last part of their chapter, the authors discuss the possibility of alternative metrics (altmetrics) to measure societal impact.

The use of econometric approaches for the measurement of research productivity, an important concept in research policy and for the wider public, is the subject of a chapter by *Cinzia Daraio*. It explains the benefits of econometric models in research assessment and shows their added value compared to more traditional bibliometric or informetric approaches. Moreover, it gives a theoretical discussion of the nature as well as the ambiguities of the concept of productivity and other key notions in research performance measurement. On the practical side, it presents a checklist for developing econometric models of research assessment.

Gunnar Sivertsen describes the development of a new type of data source for science studies. Institutional and national current research information systems (CRIS) are used to standardize and facilitate research output reporting and research administration. With their high standard of coverage, quality, and standardization, CRIS systems also have the potential to be used as data sources for science studies. Basic requirements are interoperability and data integration at the institutional and national levels. The chapter focuses on challenges and solutions to the development of internationally integrated CRIS. Challenges and possible solutions reaching far beyond the technical are described from the international level to a concrete national example. The authors also show that internationally integrated CRIS can be used for science studies.

Part D: New Indicators for Research Assessment

Indicators for academic outputs derived from social media, such as Twitter, are sometimes known as altmetrics. These are typically quicker to appear than citation counts but are not subject to peer review. These properties make social media indicators fundamentally different from citation-based indicators and there is uncertainty about how they should be used. *Paul Wouters*, *Zohreh Zahedi*, and *Rodrigo Costas* propose principles and conceptual frameworks for using social media data effectively and responsibly in research evaluation contexts. Their chapter gives practical advice as well as theory-based arguments and applies to current as well as future social media indicators.

Monographs, edited books, and book chapters are central to areas of the social sciences and humanities, and can sometimes be important outputs for other researchers. *Alesia Zuccala* and *Nicolas Robinson Garcia* review studies assessing the value of scholarly books. They show that important contributions have been made by four different expert communities, which they define as monitors, subject classifiers, indexers, and indicator constructionists. This unique perspective helps to clarify the advances that need to be made if this relatively under-researched area is to mature as a standard part of the bibliometric landscape.

The pioneering microblog site Twitter is widely used by academics to post about academic publications, such as by announcing the journal articles that they are reading. Counts of tweets about academic outputs are often described as altmetrics. *Stefanie Haustein* reviews studies about the value and interpretation of Twitter altmetrics in many fields. In addition, she uses an analysis of 24 million tweets about scholarly documents to give a detailed exploration of the context of Twitter altmetrics. This information includes the types of documents tweeted about and different types of tweeting patterns, including the problem of nonhuman tweeters.

The social reference manager Mendeley can be used for evidence of the impact of academic publications by counting the number of Mendeley users that have registered them in their personal libraries. *Ehsan Mohammadi* and *Mike Thelwall* discuss how this information is an indicator of the scholarly readership of scholarly outputs and gives citation-like impact evidence. They argue that Mendeley provides earlier impact evidence than citations and its readership data is therefore useful for research evaluations where timely impact data is important. Mendeley is also useful for the background information that it gives about readers, including their job, subject area, and national base.

Any empirical use of web indicators involves gathering data from the web at some stage. This is not as straightforward as downloading citation data from bibliometric or patent databases. *Judit Bar-Ilan* gives a historical overview of methods to gather informetric data from the web, including the main problems and proposed solutions. She shows that researchers have often had to use imperfect methods, such as queries in commercial search engines, to gather their data and this can give misleading results. She emphases that it is important for those collecting data to devote time to data cleansing and other techniques that will help to produce the most accurate and reliable information.

Although scholarly indicators derived from the social web have attracted more attention than indicators derived from the web, the latter seem to be more effective at providing evidence of nonscholarly impacts. *Kayvan Kousha* demonstrates that online data gathering, often through clever standardized queries in commercial search engines, can give indicators of educational, health, informational, general, and other impacts. The methods sometimes take advantage of individual important websites, such as Wikipedia, and sometimes search for evidence from a large part of the web. Kousha's chapter discusses the methods used to generate a range of web-based indicators and reviews evidence of their limitations and value.

Scholarly articles are usually available in electronic form and sometimes only in digital versions. Before an article can be cited, it must be accessed and this may well involve downloading it from a publisher website or a digital repository. Data from such sources may therefore give earlier evidence of the academic impact of publications and perhaps also evidence of interest from nonpublishing audiences. *Edwin A. Henneken* and *Michael J. Kurtz* demonstrate how analyzing the log files of a digital repository can give new types of detailed information about how academic research is accessed. They illustrate this with a detailed analysis of clickstream data from the Astrophysics Data System of peer reviewed and other publications about astronomy and physics.

Although most research evaluations focus mainly on journal articles and perhaps also books, a range of other types of activity and output are important to the missions of academia. It is important that the contributions of scholars producing nonstandard outputs are recognized and one way of achieving this is by generating impact indicators to support qualitative claims for their value or impact. *Mike Thelwall* introduces online indicators for different types of scholarly output that take advantage of easily available online quantitative data, such as view or download counts published online or available for automatic harvesting. Whilst interpreting the numerical information is complex due to the variety of different goals and audiences for ostensibly similar outputs, the survey shows that useful data is often available, although always with limitations.

Part E: Advancement of Methodology for Patent Analysis

The present handbook covers both science and technology indicators, as they are closely intertwined. Science is primarily linked to publication indicators, and technology to patent indicators. Nevertheless, the logic behind publication and patent statistics are quite different, so that multiple chapters are needed to explain different aspects of patent statistics. Furthermore, in addition to patents, trademarks and standards are also used as technology indicators. These alternative indicators are discussed in two chapters.

A general challenge of patent analysis is the growing number of patent applications, so that it is increasingly difficult to identify all relevant documents referring to a specific topic. For this purpose, *Carson K. Leung*, *Wookey Lee*, and *Justin Jongsu Song* present an advanced text-based retrieval system and compare three different retrieval algorithms. With their approach, patent documents that are relevant to keyword terms in a user query can be retrieved efficiently without returning many irrelevant patent documents.

A specific characteristic of patents in comparison to publications is that the text is not written by the inventors, but by patent attorneys. The latter have a decisive influence on the successful process of a patent application. *Rainer Frietsch* and *Peter Neuhäusler* analyze the differences between experienced and less experienced attorneys in more detail.

A further peculiarity of patents is that most include images of the invention. On this basis, images can be used for patent retrieval as well. *Ilias Gialampoukidis*, *Anastasia Moumtzidou*, *Stefanos Vrochidis*, and *Ioannis Kompatsiaris* describe different computer-based approaches and illustrate them with examples.

Ulrich Schmoch and *Mosahid Khan* deal with new methodological issues of retrieval for patent indicators linked to the change of the patent system in the last 20 years and the new ways to access patent data. It describes international flows of patent applications between the US, Europe, and Southeast Asia and illustrates methods for an appropriate cross-country comparison. A central topic of this chapter is the implications of the frequently used Patent Cooperation Treaty (PCT) route of patent applications on the conception of search strategies and the interpretation of search results. Furthermore, the possibilities of search with the new international Cooperative Patent Classification (CPC) are explained. In addition, the patenting activities of very large companies and the patent value are discussed.

For knowledge-based technologies, scientific and technological activities are performed in parallel and influence each other. Here it is useful to identify similar patents and publications. *Tom Magerman* and *Bart Van Looy* present linguistic text mining approaches to identify similarities and illustrate them with examples. They discuss the advantages and disadvantages of different retrieval methods.

Contributions in this handbook show that more powerful computer systems have increased the power of text mining in science and technology analysis. *Samira Ranaei*, *Arho Suominen*, *Alan Porter*, and *Tuomo Kässi* give a broad literature review of the most relevant approaches and show by examples their usefulness. For instance, text mining classification analysis of patents can lead to additional results. It is insightful to compare the methods of other chapters on text mining in this handbook to the assessment of the authors of this chapter.

Other contributions in this handbook also illustrate the potential of text mining. However, the quality of results are influenced by the quality of the texts analyzed. Thus, the yield of text-based retrieval at some patent offices is higher than at others depending on the legal requirements for technical disclosure of the patent abstracts. For instance, the text quality at the US Patent and Trademark Office is commendable.

An advantage of patents compared to scientific publications is their detailed classifications, so that in many cases a precise definition of a topic is feasible. However, it is possible that similar items are classified in different parts of the classification, so that it is difficult to identify all documents relevant to a topic. *Andrea Bonaccorsi*, *Gualtiero Fantoni*, *Riccardo Apreda*, and *Donata Gabelloni* suggest a functional classification system for patents which supports new types of patent searches based on functional dictionaries. Again, the approach is based on advanced text mining. The authors present some examples of contexts for which their approach is useful.

Part F: Patent System, Patents, and Economics

A widespread misconception is that software inventions can be patented only in the USA, but not in Europe. The chapter of *Peter Neuhäusler* and *Rainer Frietsch* shows that in many cases software can also be patented in Europe and that the share within all patent applications is steadily increasing. In addition, the introduction of the subclass G06Q (data processing systems or methods, specially adapted for administrative, commercial, financial, managerial, supervisory, or forecasting purposes) in 2006 confirms that the attitude of the European Patent Office about software has substantially changed.

Patents and trademarks are competitive tools for research-intensive technologies. *Sandro Mendonça*, *Ulrich Schmoch*, and *Peter Neuhäusler* show with the example of enterprises from the EU Industrial R&D Scoreboard that trademarks, product and service marks have become increasingly important, especially in the case of service marks. They illustrate with examples the fact that the strategies of enterprises for patents and trademarks vary considerably between sectors and sometimes even within sectors. They argue that trademarks should be considered in parallel to patents whenever possible.

In most countries, the annual number of patent applications is stable and changes only in the long term. A new phenomenon for threshold countries, such as South Korea and China, is the tremendous increase in the annual number of patent applications within a decade. *Chan-Yuan Wong* and *Hon-Ngen Fung* analyze South Korea and China, highlighting parallel increases in scientific activities. These types of explorations will be important for understanding the worldwide landscape of science and technology and the emergence of a new regime in international trade.

The final contribution to the technology section of this handbook is concerned with standards. It shows that new technologies are not sufficient for market success, but that the development of standards is a further decisive step. *Knut Blind* explains why standards are important for technology and that they can be used as indicators for describing supplementary aspects of technological performance.

References

- A. F. J. van Raan (Ed.): *Handbook of Quantitative Studies of Science and Technology* (North Holland, Amsterdam 1988)
- H. Moed, W. Glänzel, U. Schmoch (Eds.): *Handbook of Quantitative Science and Technology Research* (Springer, Dordrecht 2004)
- C. Daraio, W. Glänzel: Grand challenges in data integration—state of the art and future perspectives: An introduction, Scientometrics **108**(1), 391–400 (2016)

Wolfgang Glänzel
Henk F. Moed
Ulrich Schmoch
Mike Thelwall

About the Editors

Wolfgang Glänzel is Director of the Center for Research and Development Monitoring (ECOOM) and Full Professor at KU Leuven, Belgium. He is also affiliated with the Department of Science Policy and Scientometrics at the Library of the Hungarian Academy of Sciences in Budapest, Hungary. Wolfgang Glänzel studied Mathematics at the Eötvös Lorand University (ELTE) in Budapest. He holds a doctorate in Mathematics from ELTE obtained in 1984 as well as a PhD in the Social Sciences which he obtained from Leiden University, The Netherlands, in 1997. He is also Guest Professor at several Universities in China and the UK.

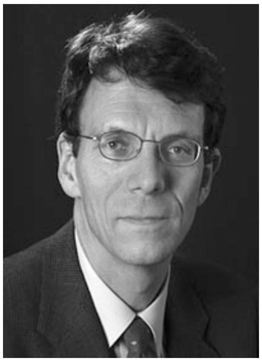
Wolfgang Glänzel's research activities comprise probability theory and mathematical statistics, quantitative science studies, and research policy. In the field of probability theory he has published on the characterization of probability distributions, in quantitative science studies and research policy his research is mainly focused on models of scientific communication, the development of scientometric indicators, the analysis of coauthorship patterns and research collaboration, the structural-cognitive mapping of science, bibliometrics-aided retrieval, and the measurement of research performance at various levels of aggregation. Wolfgang Glänzel has published numerous journal articles, proceedings, and book chapters and coauthored/coedited several books. He was Alexander von Humboldt Fellow for two years in Germany. He has published in probability theory and mathematical statistics, computer science and various topics in quantitative science studies, scientometrics, and informetrics.

In 1999, Professor Glänzel received the international *Derek deSolla Price Award* for outstanding contributions to the quantitative studies of science. He is Secretary-Treasurer of the *International Society for Scientometrics and Informetrics* (ISSI) and Editor-in-Chief of the Society's Newsletter. Wolfgang Glänzel has been Editor-in-Chief of the international journal *Scientometrics* since 2014 and Academic Editor of the multidisciplinary OA journal PLoS ONE since 2013.

Henk F. Moed holds a PhD degree in Science Studies and was a senior staff member and Full Professor in Research Assessment Methodologies at the Centre for Science and Technology Studies at Leiden University, The Netherlands between 1981 and 2010. He was a Senior Scientific Advisor at Elsevier, Amsterdam, from 2010–2014.

Henk Moed has been active in numerous research topics, including: the creation of bibliometric databases; the development and use of bibliometric indicators; studies of the effects of *Open Access*; bibliometric studies of international scientific migration and collaboration; comparisons of Web of Science, Scopus and Google Scholar; multidimensional assessment of research impact; the potential of altmetrics; ontology-based bibliometric data management.

Dr. Moed published over 100 research articles in international, peer-reviewed journals, was Program Chair of numerous international conferences in the field, and is an Editorial Board Member of several journals in his field. He was a winner of the *Derek deSolla Price Award* for outstanding contributions to the quantitative studies of science in 1999. He is currently an independent scientific advisor, and Visiting Professor at the Sapienza University of Rome and University of Granada.



Ulrich Schmoch holds a Doctoral Degree in Social and Political Sciences (University of Hannover, Germany), a Diploma of Mechanical Engineering (University of Hannover) with modules at the University of Grenoble, France, and a State Doctorate in Sociology of Science and Technology (Karlsruhe Institute of Technology KIT, Germany). From 1983–1985, he was employed in the office of a patent attorney.

Since 1986, Professor Schmoch has worked at the Fraunhofer Institute for Systems and Innovation Research (ISI), Karlsruhe, Germany, and performed and coordinated about 200 projects in the areas of innovation indicators, patent analysis, bibliometrics, systems of innovation, technology transfer, science management, innovation in threshold countries, innovations in the service sector, science-based technologies, and organization and mission of universities in an international comparison.

Between 1991–2002, Ulrich Schmoch was Deputy Head and from 2002–2004, Head of the Department *Technology Assessment and Innovation Strategies*. In 2011, he taught innovation indicators and patent statistics at the University of Stellenbosch, South Africa. Between February 2012 and September 2014 he taught as Stand-In Professor in Science Management at the German University of Public Administration in Speyer. Since 2015, he has once again been working at the Fraunhofer ISI as a Senior Researcher.

Mike Thelwall is a Professor of Data Science and leads the Statistical Cybermetrics Research Group at the University of Wolverhampton, UK, which he joined in 1989. He is also Docent at the Department of Information Studies at Åbo Akademi University in Turku, Finland, and a Research Associate at the Oxford Internet Institute. He obtained a PhD in Pure Mathematics from the University of Lancaster, UK.

Professor Thelwall has developed tools and methods for gathering and analyzing webometric and altmetric alternative indicators. His publications include 300 refereed journal articles, 30 book chapters, and 3 books.

Mike Thelwall is an Associate Editor of the Journal of the Association for Information Science and Technology and sits on four other editorial boards. He was a member of the UK's independent review of the role of metrics in research assessment (2014–2015) and is a member of the UK Forum for Responsible Research Metrics.

List of Authors

Riccardo Apreda
Erre Quadro s.r.l.
Pisa, Italy

Judit Bar-Ilan
Information Science
Bar-Ilan University
Ramat Gan, Israel

Knut Blind
Institute of Technology and Management
Technische Universität Berlin
Berlin, Germany

Andrea Bonaccorsi
DESTEC
University of Pisa
Pisa, Italy

Katy Börner
School of Informatics, Computing,
and Engineering
Indiana University
Bloomington, IN, USA

Lutz Bornmann
Division for Science and Innovation Studies
Administrative Headquarters of the
Max Planck Society
Munich, Germany

Kevin W. Boyack
SciTech Strategies, Inc.
Albuquerque, NM, USA

Linda Butler
School of Politics and International Relations
The Australian National University
Acton, Australia

Guillaume Cabanac
Computer Science Department
University of Toulouse
Toulouse, France

Martine Cadot
Knowledge and Language Management
LORIA (Laboratory in Computer Science and its
Applications)
Villers-lès-Nancy, France

Manuel J. Cobo
Department of Computer Science and Engineering
University of Cádiz
Algeciras, Spain

Rodrigo Costas
Centre for Science and Technology Studies (CWTS)
Leiden University
Leiden, The Netherlands

Cinzia Daraio
Department of Computer, Control,
and Management Engineering
Sapienza University of Rome
Rome, Italy

Koenraad Debackere
ECOOM and Faculty of Economics and Business
KU Leuven
Leuven, Belgium

Emilio Delgado López-Cózar
Información y Documentación
Universidad de Granada
Granada, Spain

Paul Donner
Research Area 2, Research System and Science
Dynamics
German Centre for Higher Education Research
and Science Studies, DZHW
Berlin, Germany

Nees Jan van Eck
Centre for Science and Technology Studies (CWTS)
Leiden University
Leiden, The Netherlands

Gualtiero Fantoni
Department of Civil and Industrial Engineering
University of Pisa
Pisa, Italy

Rainer Frietsch
Competence Center Policy – Industry – Innovation
Fraunhofer Institute for Systems and Innovation
Research ISI
Karlsruhe, Germany

Hon-Ngen Fung
Post-graduate and Research Centre
KDU University College
Shah Alam, Malaysia

Donata Gabelloni
Erre Quadro s.r.l.
Pisa, Italy

Ilias Gialampoukidis
Information Technologies Institute
Centre for Research and Technology Hellas
Thermi, Greece

Gali Halevi
Icahn School of Medicine at Mount Sinai
New York, NY, USA

Robin Haunschild
Central Information Service
Max Planck Institute for Solid State Research
Stuttgart, Germany

Stefanie Haustein
School of Information Studies
University of Ottawa
Ottawa, Canada

Edwin A. Henneken
High Energy Physics
Harvard-Smithsonian Center for Astrophysics
Cambridge, MA, USA

Enrique Herrera-Viedma
Department of Computer Science and Artificial Intelligence
University of Granada
Granada, Spain

Sybille Hinze
Research System and Science Dynamics
German Center for Higher Education Research and Science Studies (DZHW)
Berlin, Germany

Xiaojun Hu
School of Basic Medical Sciences
Zhejiang University
Hangzhou, China

Inga Ivanova
Institute for Statistical Studies and Economics of Knowledge
National Research University Higher School of Economics (NRU HSE)
Moscow, Russian Federation

Tuomo Kässi
School of Engineering Science, Industrial Engineering and Management
Lappeenranta University of Technology (LUT)
Lappeenranta, Finland

Mosahid Khan
World Intellectual Property Organization (WIPO)
Geneva, Switzerland

Richard Klavans
SciTech Strategies, Inc.
Wayne, PA, USA

Ioannis Kompatsiaris
Information Technologies Institute
Centre for Research and Technology Hellas
Thermi, Greece

Kayvan Kousha
Statistical Cybermetrics Research Group
University of Wolverhampton
Wolverhampton, UK

Michael J. Kurtz
Optical and Infrared Astronomy
Harvard-Smithsonian Center for Astrophysics
Cambridge, MA, USA

Vincent Larivière
École de bibliothéconomie et des sciences de l'information
Université de Montréal
Montréal, Canada

Wookey Lee
Industrial Engineering
Inha University
Incheon, South Korea

Alain Lelu
ELLIADD (retired)
Université de Franche-Comté
Besançon, France

Maurizio Lenzerini
Department of Computer, Control, and Management Engineering
Sapienza University of Rome
Rome, Italy

Jacqueline Leta
Programa de Educação, Gestão e Difusão / Instituto de Bioquímica Médica
Universidade Federal do Rio de Janeiro
Rio de Janeiro, Brazil

Carson Leung
Department of Computer Science
University of Manitoba
Winnipeg, Canada

Grant Lewison
Division of Cancer Studies
King's College London
London, UK

Loet Leydesdorff
Amsterdam School of Communication Research (ASCoR)
University of Amsterdam
Amsterdam, The Netherlands

Antonio G. López-Herrera
Department of Computer Science and Artificial Intelligence
University of Granada
Granada, Spain

Roberto Mario Lovón Canchumani
Graduate Program in Information Science/IBICT-UFRJ
Federal University of Rio the Janeiro
Rio de Janeiro, Brazil

Zheng Ma
Institute of Scientific and Technical Information of China (ISTIC)
Beijing, China

Tom Magerman
Faculty of Economics and Business
KU Leuven
Leuven, Belgium

Alberto Martín-Martín
Información y Documentación
Universidad de Granada
Granada, Spain

Ian McAllister
School of Politics and International Relations
The Australian National University
Acton, Australia

Sandro Mendonça
ISCTE University Institute of Lisbon
Lisboa, Portugal

Martin Meyer
Kent Business School
University of Kent
Canterbury, UK

Staša Milojević
School of Informatics, Computing, and Engineering
Indiana University
Bloomington, IN, USA

Ehsan Mohammadi
School of Library and Information Science
University of South Carolina
Columbia, SC, USA

Jose A. Moral-Munoz
Department of Nursing and Physiotherapy
University of Cádiz
Cádiz, Spain

Anastasia Moumtzidou
Information Technologies Institute
Centre for Research and Technology Hellas
Thermi, Greece

Peter Neuhäusler
Competence Center Policy – Industry – Innovation
Fraunhofer Institute for Systems and Innovation Research ISI
Karlsruhe, Germany

Raymundo das Neves Machado
Departamento de Fundamentos e Processos Informacionais, Instituto de Ciência da Informação
Universidade Federal da Bahia
Salvador, Brazil

Enrique Orduña-Malea
Audiovisual Communication, Documentation, and History of Art
Universitat Politècnica de València
Valencia, Spain

Elena Pallari
Centre for Implementation Science, Health Service & Population Research Department
King's College London
London, UK

Alan Porter
Search Technology, Inc.
Norcross, GA, USA

Anthony van Raan
Centre for Science and Technology Studies (CWTS)
Leiden University
Leiden, The Netherlands

Samira Ranaei
School of Engineering Science, Industrial Engineering and Management
Lappeenranta University of Technology (LUT)
Lappeenranta, Finland

Nicolas Robinson-García
School of Public Policy
Georgia Institute of Technology
Atlanta, GA, USA

Ronald Rousseau
ECOOM
KU Leuven
Leuven, Belgium

András Schubert
Library and Information Center
Hungarian Academy of Sciences
Budapest, Hungary

Gábor Schubert
University Library
Stockholm University
Stockholm, Sweden

Gunnar Sivertsen
NIFU – Nordic Institute for Studies of Innovation, Research and Education
Oslo, Norway

Justin Jongsu Song
Industrial Engineering
Inha University
Incheon, South Korea

Cassidy R. Sugimoto
School of Informatics and Computing
Indiana University Bloomington
Bloomington, IN, USA

Arho Suominen
VTT Technical Research Centre of Finland
Espoo, Finland

Bart Thijs
ECOOM
KU Leuven
Leuven, Belgium

Daniel Torres-Salinas
EC3metrics SL y Medialab-UGR
Universidad de Granada
Granada, Spain

Bart Van Looy
Faculty of Economics and Business
KU Leuven
Leuven, Belgium

Stefanos Vrochidis
Information Technologies Institute
Centre for Research and Technology Hellas
Thermi, Greece

Ludo Waltman
Centre for Science and Technology Studies (CWTS)
Leiden University
Leiden, The Netherlands

Lorna Wildgaard
Researcher Support
The Royal Library
Copenhagen, Denmark

Chan-Yuan Wong
Institute of Technology Management
National Tsing Hua University
Hsinchu City, Taiwan

Paul Wouters
Centre for Science and Technology Studies (CWTS)
Leiden University
Leiden, The Netherlands

Zohreh Zahedi
Centre for Science and Technology Studies (CWTS)
Leiden University
Leiden, The Netherlands

Lin Zhang
School of Information Management
Wuhan University
Wuhan, China

Michel Zitt
Lereco Lab (retired)
National Institute for Agronomic Research (INRA), Dept SAE2
Nantes, France

Alesia Zuccala
Department of Information Studies
University of Copenhagen
Copenhagen, Denmark

Contents

Part C Science Systems and Research Policy

Part F Patent System, Patents and Economics

List of Abbreviations

5G	fifth-generation wireless system

A

ABM	agent-based model
ADR	accumulation for dynamic ranking
ADS	astrophysics data system
AHDH	adaptive hierarchical density histogram
AHP	analytic hierarchy process
AIA	America Invents Act
AIDS	acquired immune deficiency syndrome
AIS	article influence score
AKM	axial *k*-means
ALIRP	author-level indicators of research production
ALM	article-level metrics
AMP	agent modeling platform
AN	attribute network
ANOVA	analysis of variance
ANT	actor–network theory
AP	advanced placement
AP	average precision
APC	article processing charges
API	application programming interface
ARC	Australian Research Council
ARI	adjusted Rand index
ASJC	All Science Journal Classification
ASM	atomic structure model
ASR	American Sociological Review
ATECO	attività economiche
AVA	autovalutazione, valutazione periodica, accreditamento
A&H	arts and humanities
A&HCI	Arts and Humanities Citation Index

B

BC	bibliographic coupling
BCCL	book classification for Chinese libraries
BDOA	Berlin Declaration on Open Access
Belspo	Belgian Federal Science Policy Office
BIS	bibliographic information system
BKCI	Book Citation Index
BKCI-S	Book Citation Index – Science
BKCI-SSH	Book Citation Index – Social Sciences & Humanities
BM25	best match 25
BMC	BioMed Central
BMJ	British Journal of Medicine
BOAI	Budapest Open Access Initiative
BOF	extraordinary research fund
BoW	bag-of-words
BRICS	Brazil, Russia, India, China and South Africa

C

CA	correspondence analysis
CAGR	compound annual growth rate
CARDI	cardiovascular disease
CAS	chemical abstracts service
CASRAI	Consortia Advancing Standards in Research Administration Information
CBA	cost–benefit analysis
CBIR	content-based image retrieval
CC	co-citation
ccTLD	country code top-level domain
CHSSCD	Chinese Humanities and Social Sciences Core Journals Database
CII	computer-implemented invention
CISE	computer and information science and engineering
CJCR	Chinese STM Citation Report
CJK	China, Japan, South Korea
CNCI	category normalized citation impact
COL	colorectal cancer
COPD	chronic obstructive pulmonary disease
COR	coronary heart disease
CoreSC	core scientific concept
COUNTER	Counting Online Usage of Networked Electronic Resources
CPC	Cooperative Patent Classification
CPCI-S	Conference Proceedings Citation Index – Science
CPCI-SSH	Conference Proceedings Citation Index – Social Science & Humanities
CPG	clinical practice guideline
CPS	comprehensive performance score
CQ	conjunctive queries
CRIS	current research information system
CRISTIN	Current Research Information System in Norway
CSCD	Chinese Science Citation Database
CSS	characteristic scores and scales
CSSCI	Chinese Social Sciences Citation Index
CSTPCD	Chinese Scientific and Technical Papers and Citations Database
CSV	comma-separated values
CWTS	Centre for Science and Technology Studies

D

DALY	disability-adjusted life years
DBSCAN	density-based spatial clustering of application with noise
DC	direct citation
DCI	data citation index
DDC	Dewey Decimal Classification System

DDR	dispersion for dynamic ranking
DEA	data envelopment analysis
DGP	data generating process
DIABE	diabetes
DL	digital library
DOAJ	Directory of Open Access Journals
DOI	digital object identifier
DORA	San Francisco Declaration on Research Assessment
DORA	Declaration on Research Assessment
DSM	distributional semantic model
DSSC	dye-sensitized solar cells

E

ECLA	European classification
EDC	extended direct citation
EEA	European Economic Area
Ei	Engineering index
EM	expectation maximization algorithm
EMNPC	equalized mean-based normalized proportion cited
EMR	Elastic MapReduce
EOAC	edge orientation autocorrelogram
EPC	European Patent Convention
ER	entity resolution
ERA	Excellence in Research for Australia
ERiC	Evaluating Research in Context
ERIH	European Reference Index for the Humanities
ESCI	Emerging Sources Citation Index
ESI	Essential Science Indicators
ET	emerging technology
ETER	European Tertiary Education Register
EU FP7/H2020	EU Framework Program/Horizon 2020
EuroCRIS	European Current Research Information Systems
EV	electric vehicle

F

FAIR	findability, accessibility, interoperability, and reusability
FBS	function–behavior–structure
FCEV	fuel cell electric vehicle
FCM	fuzzy C-means method
FDI	foreign direct investment
FECR	field expected citation rate
FN	false negative
FoR	fields of research
FoS	fields of science and technology
FP	false positive
FPC	functional patent classification
FRBR	Functional Requirements for Bibliographic Records
FSS	fractional scientific strength
FTE	full time equivalent
FUSE	Foresight and Understanding from Scientific Exposition
FWCI	field-weighted citation impact
FWO	fund for scientific research Flanders

G

GAV	global-as-view
GDP	gross domestic product
GIF	Global Impact Factor
GLM	generalized linear model
GML	Graph Modeling Language
GMM	generalized method of moment
GPL	general public license
GPU	graphics processing unit
GRID	Global Research Identifier Database
GS	Google Scholar
GSC	Google Scholar Citations
GSM	Google Scholar Metrics
GtR	Gateway to Research
GUESS	Graph Exploration system
GUI	graphical user interface

H

HCE	hit count estimate
HCR	highly cited researchers
HHI	Herfindahl–Hirschman index
HIV	human immunodeficiency virus
HTML	hypertext markup language
HWWS	handwashing with soap

I

ICA	independent component analysis
ICE	internal combustion engine
ICEE	Indicator of Quality for Publishers according to Experts
ICT	information and communication technology
ICV	Index Copernicus metric value
IDF	inverted document frequency
IDR	interdisciplinary research
IF	impact factor
IFBSCP	Impact Factor Biased Self-Citation Practices
IM	intermediary
INPADOC	International Patent Documentation
IoT	Internet of Things
IP	intellectual property
IPC	International Patent Classification
IPR	intellectual property rights
IR	information retrieval
ISBN	International Standard Book Number
ISI	Institute for Scientific Information
ISNI	International Standard Name Identifier
ISO	International Organization for Standardization
IT	information technology

J

J-STAGE	Japan Science and Technology Information Aggregator
JASIST	Journal of the Association for Information Science and Technology
JASSS	Journal of Artificial Societies and Social Simulation
JCR	Journal Citation Reports
JID	Journal IDentification
JIF	Journal Impact Factor

K

KET	key enabling technology
KIS	knowledge-intensive services
KML	keyhole markup language
KTN	knowledge transfer network
KWD	keyword weight distribution

L

LBC	Library-Bibliographical Classification
LBD	literature-based discovery
LBT	link-bridged topic model
LC	longitudinal coupling
LCC	Library of Congress Classification
LDA	latent Dirichlet allocation
LIS	library and information science
LMI	lead market initiative
LOWESS	locally weighted scatterplot smoothing
LSA	latent semantic analysis
LSH	locality-sensitive hashing
LSI	latent semantic indexing
LUBM	Lehigh University Benchmark

M

MA	Microsoft Academic
MAG	Microsoft Academic Graph
MBS	mean blog score
MCS	mean citation score
MDS	multidimensional scaling
MECR	mean expected citation rate
MENTH	mental disorder
MeSH	Medical Subject Headings
MHT	medium-high-technology
MNCS	mean normalized citation score
MNE	multinational enterprise
MNLCS	Mean Normalized Log-transformed Citation Score
MOCR	mean observed citation rate
MRI	magentic resonance imaging
MRS	mean readership score
MSC	mathematics subject classification
MTS	mean Twitter score
M&A	mergers and acquisitions

N

NACE	Statistical Classification of Economic Activities in the European Community
NB	naive Bayes
NBIC	nanotechnology, biotechnology, information technology, and cognitive science
NCD	normal compression distance
NEST	newly emerging science and technology
NGD	normalized Google distance
NGO	nongovernmental organization
NLP	natural language processing
NLTK	Natural Language Toolkit
NMF	non-negative matrix factorization
NMI	normalized mutual information
NPD	new product development
NPR	nonpatent reference
NVL	Nationale VersorgungsLeitlinien
nwD	neighbor-weighted degree
NWE	neural word embeddings

O

OA	open access
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
OAMJ	open access mega-journal
OBDA	ontology-based data access
OBDI	ontology-based data integration
OBDM	ontology-based data management
OCR	optical character recognition
OCR	over-citation ratio
OJS	open journal systems
OLS	ordinary least squares
OOB	out-of-bag
OP	open factor
ORCID	Open Researcher and Contributor ID
OWL	ontology web language

P

PACS	physics and astronomy classification scheme
PAN	publication-attribute network
PATSTAT	EPO Worldwide Patent Statistical Database
PCA	principal component analysis
PCT	Patent Cooperation Treaty
PEM	proton-exchange membrane
PFI	Pact for Research and Innovation
PHI	PartnersHIp Ability Index
PI	principal investigator
PISA	Program for International Student Assessment
PLoS	Public Library of Science
pLSA	probabilistic latent semantic analysis
PLSI	probabilistic latent semantic indexing
PN	publication network

PNG	portable network graphics
POS	part of speech
PP(top 10%)	proportion of top 10% publications
ppp	purchasing power parity
PPV	positive predictive value
PRI	public research institution
PRO	public research organization

Q

QMS	quality management system

R

RAE	Research Assessment Exercise
RCD	Research Core Data Set
RCI	relative citation impact
RDBMS	relational database management system
RDI	research and development intensity
REF	Research Excellence Framework
RePEc	Research Papers in Economics
REPP	research embedment and performance profile
RESSH	Research Evaluation in the Social Sciences and Humanities
RF	random forest
RG	ResearchGate
RIS	research information system
RJ	Riksbankens Jubileumsfond
RO	receiving office
RoW	rest of the world
RPYS	referenced publication years spectroscopy
RQ	research quantum
RQF	Research Quality Framework
RSI	relative specialization index
R&D	research and development
R&I	research and innovation

S

SAO	subject–action–object
SAS	Statistical Analysis System
SBIR	Small Business Innovation Research
SC	subject category
SCI	Science Citation Index
Sci^2	Science of Science
SCI-WoS	Science Citation Index of the Web of Science
SCIE	Science Citation Index Expanded
SciMAT	science mapping analysis software tool
SciSTIP	Scientometrics and Science, Technology and Innovation Policy
SERP	search engine results page
SFA	stochastic frontier analysis
SIAMPI	Social Impact Assessment Methods for research and funding instruments
SIFT	scale invariant features transform
SJR	SCImago Journal Rank
SLM	smart local moving
SLMA	smart local moving algorithm
SMART	specific, measurable, accessible, relevant, and traceable
SME	small and medium-sized enterprise
SMOTE	synthetic minority oversampling technique
SNA	system of national account
SNIP	source normalized impact per paper
SNLG	Sistema Nazionale per le Linee Guida
SOM	self-organizing map
SOOS	steunpunt o&o statistieken
SOSP	Science of Science Policy
SPSS	statistical package for the social sciences
SQL	structured query language
SS	social sciences
SSCI	Social Sciences Citation Index
SSH	social sciences and humanities
ST	single terms
ST	science and technology
STEM	science, technology, engineering, and mathematics
STI	science, technology, and innovation
STICCI	software tool for improving and converting citation indices
STIP	science, technology, and innovation policy
STM	scientific, technical and medical
STN	Science Technology Network
STS	SciTech Strategies
STS	science and technology studies
SURF	speeded-up robust features
SVD	singular value decomposition
SVG	scalable vector graphics
SVM	support vector machine
S&T	science and technology

T

TA	threshold algorithm
TARL	topics, aging, and recursive linking
tAS	triple adjacent segment
TBT	technical barrier to trade
TED	technology, entertainment, design
TF	term frequency
TF	technological forecasting
TF-ICF	term frequency-inverse corpus frequency
tf-idf	term frequency-inverse document frequency
TF-IPCF	term frequency-inverse patent category frequency
TH	triple helix
THE	Times Higher Education
TIMSS	Trends in International Mathematics and Science Study

TLD	top-level domain
TN	true negative
TOA	technological opportunities analysis
TOPSIS	technique for order of preference by similarity to ideal solution
TP	true positive
TPR	true positive rate
TRIPS	Trade Related Aspects of Intellectual Property Rights
TRIZ	theory of inventive problem solving
TRM	technology road mapping
TTIP	Transatlantic Trade and Investment Partnership

U

UDC	universal decimal classification
UIF	Universal Impact Factor
UML	unified modeling language
UNA	unique name assumption
UoA	unit of assessment
URL	uniform resource locator
USPC	US patent classification

V

VABB-SHW	Vlaams Academisch Bibliografisch Bestand voor de Sociale en Humane Wetenschappen
VBA	visual basic application
VEM	variational expectation maximization
VQR	Valutazione della Qualità della Ricerca—Evaluation of Research Quality
VSM	vector space model

W

WIF	web impact factor
WIRe	web impact report
WoS	Web of Science
WPI	World Patents Index
WSDM	Web Search & Data Mining

X

XML	extensible markup language
XRAC	cross-reference art collection