



科研评价中

的引文分析

CITATION ANALYSIS IN RESEARCH EVALUATION

亨克·F. 莫德(荷) 著

佟贺丰 等 译

王 炼 等 译校

庞景安 审定

科研评价中的引文分析

亨克·F. 莫德(荷) 著

佟贺丰 王 炼 曹 燕

黄军英 蔡荣海 刘润生

于 洁 廖 凤 吴善略 译

王 炼 佟贺丰 徐 峰 校

庞景安 审定

科学技术文献出版社

Scientific and Technical Documents Publishing House

北 京

图书在版编目(CIP)数据

科研评价中的引文分析/佟贺丰等译. -北京:科学技术文献出版社,2010.5
ISBN 978-7-5023-6666-7

I. ①科… II. ①佟… III. ①科学研究-评估-引文索引-分析 IV. ①G311

中国版本图书馆 CIP 数据核字(2010)第 081526 号

出 版 者 科学技术文献出版社
地 址 北京市复兴路 15 号(中央电视台西侧)/100038
图书编务部电话 (010)58882938,58882087(传真)
图书发行部电话 (010)58882866(传真)
邮 购 部 电 话 (010)58882873
网 址 <http://www.stdph.com>
E-mail:stdph@istic.ac.cn
策 划 编 辑 周国臻
责 任 编 辑 周国臻
特 约 编 辑 李琴燕
责 任 校 对 唐 炜
责 任 出 版 王杰馨
发 行 者 科学技术文献出版社发行 全国各地新华书店经销
印 刷 者 北京时尚印佳彩色印刷有限公司
版 (印) 次 2010 年 5 月第 1 版第 1 次印刷
开 本 787×1092 16 开
字 数 339 千
印 张 15
印 数 1~1000 册
定 价 38.00 元

© 版权所有 违法必究

购买本社图书,凡字迹不清、缺页、倒页、脱页者,本社发行部负责调换。

(京)新登字 130 号

内 容 简 介

本书主要围绕“在科研评价中如何正确地使用引文分析这样一种工具?”这一问题展开。书中纠正了一些有关引文分析的常见误解,既揭示了引文分析对学术文献进行量化、文献计量学分析的巨大作用,能够帮助文献计量评估与研究人员更深层次地理解学术活动和绩效,突出其政策关联性;也指出了引文分析在科研评价中的限制,并对可能存在的滥用提出了警告;提出了将引文分析作为一种科研评价工具的正确使用的标准。

本书重点介绍的是如何利用美国科学信息研究所(ISI)出版的科学引文索引(SCI)数据有效地开展引文分析评价研究以及制定相关的政策。

本书适合从事科研绩效评估的科学家和学者、科研政策的决策制定者、高等院校与科研机构的科研管理人员,以及图书馆学、信息科学和科学计量学领域的教师和学生阅读参考。

译者序

引文分析作为一种应用广泛的文献计量学研究方法,早在 20 世纪 20 年代就已出现。但直到 60 年代初尤金·加菲尔德(Eugene Garfield)创建了“科学引文索引”(SCI, Science Citation Index),引文分析才真正得到普及。引文分析利用引文数据构造并应用了一系列反映学术成果的“影响力”或“质量”的指标,这些指标被应用于学术交流研究和科研绩效评价。

尤金·加菲尔德指出:“引文分析是科学管理的有力工具,通过引文分析可以了解科学史和科学结构,同样引文分析也可以用来评述科学期刊和科学家。”政府和研究机构需要对科研产出做系统化的评估,以优化研究资源分配、重新定向研究支持方向、改革研究组织、重新构造特殊领域的研究或者提高研究产出。现在,很多国家都把引文分析作为不同层次科研评价的重要手段和工具。引文分析正式被我国接受和运用是在 1980 年前后,国内机构和学者对引文分析的理论和实践进行了诸多研究与尝试,我国政府决策部门也逐渐把引文分析作为决策和政策指定的重要参考工具。

引文分析从诞生之日起,就伴随着很多争论——通过对引文的分析,能否达到事先预设的目标。特别是在引文分析成为科研评价的工具、科技资源配置的决策基础时,争论就更加激烈,因为这关系到一个国家、机构、团队或个人的科研发展前景。

《科研评价中的引文分析》就针对这个疑问做了详细的理论与实证研究。本书主要围绕“在科研评价中如何正确地使用引文分析这样一种工具?”这一问题展开。书中纠正了一些有关引文分析的常见误解,既揭示了引文分析对学术文献进行定量化、文献计量学分析的巨大作用,能够帮助更深层次地理解学术活动和绩效,并突出其政策关联性;也指出了引文分析在科研评价中的限制,并对可能存在的滥用提出了警告。本书提出了将引文分析作为一种科研评价工具的正确使用的标准。

将引文分析正确地应用于科研评价之中有较高的能力要求,并不是任何一个能使用 ISI 引文索引的人都能够轻易完成的。本书的研究表明,正确的引文分析需要详细的技术知识和理论理解。如果对引文分析的了解不够深入,就会在科研评价中闹笑话。比如国内一些科研机构规定,在对科研人员发表论文进行奖励时,以 1 000 元为基数,再乘以文章发表期刊的影响因子。这样发表在高影响因子期刊上的论文作者就能够得到足够高的奖金。可以说本意是好的,但这种做法完全是对引文分析方法的不了解。首先并不存在引文分析衡量的质量就是实际的质量的说法;其次不同学科领域间的期刊影响因子不具有可比性,比如数学领域期刊的影响因子一般就大大低于生物化学和分子生物学期刊。估计数学系的教师如果知道这个规律,肯定会在心里叫屈。

然而,本书并非是提供如何进行引文分析的详细技术手册。相反,本书的重点是引文分析

的主要框架、基本原则和假设,以及引文分析在各学术领域的应用和限制。本书主要介绍了如何评价学者通过研究成果对正确的学术知识做出的贡献。本书讨论了对单个学者、研究小组、部门和机构、学术期刊和国家学术体系的科研绩效的评价,同时分析了全球科学和学术的一般特征。本书的目的是,通过系统讨论学者和决策者有关引文分析价值的多种观点、纠正并指明引文分析的长处和短处(特别是相对于同行评议),为学术界的成员和制定研究政策的官员提供有关引文分析的基本技术方面的有用信息,包括引文分析衡量什么、如何将其合适地应用于科研评价和政策制定过程当中。

本书的作者亨克·莫德(Henk F. Moed)教授,任教于荷兰莱顿大学(Leiden University)科技研究中心。莫德教授是文献计量学领域的资深专家,1999 年度普赖斯奖得主。他的这本《科研评价中的引文分析》2005 年由 Springer 公司出版。现由中国科学技术信息研究所引进该书版权并翻译出版。在此感谢莫德教授和 Springer 出版公司的工作。希望《科研评价中的引文分析》中文版的面世,可以为我国的科研评价工作提供一些指导。同时,希望本书的出版,可以让科研评价的参与人员能够更深入地了解引文分析的本质、引文分析的指标是如何构建和计算的、有关衡量对象的各种理论观点,以及引文分析在科研评价中的优缺点;可以为学术界的成员和制定研究政策的官员提供有关引文分析的基本技术方面的有用信息,特别是如何将其恰当地应用于科研评价和政策制定过程当中。

全书的翻译由庞景安和佟贺丰进行了具体的组织安排,因为本书的篇幅较长,有多位研究人员参与了具体的翻译工作,具体的分工如下:前言、概述、第 1 章和第 2 章由王炼翻译;第 3 章由曹燕翻译;第 4 章和第 5 章由黄军英翻译;第 6~8 章由蔡荣海翻译;第 9~14 章由刘润生翻译;第 15~17 章、第 26 章和第 27 章由廖凤翻译;第 18 章和第 19 章由于洁翻译;第 20 章由吴善略翻译;第 21~25 章由佟贺丰翻译。翻译完成后由王炼、佟贺丰和徐峰进行了仔细的校对、修改。全书最后由庞景安研究员审定。科学技术文献出版社周国臻做了大量的译校与编辑工作。因为时间和水平有限,翻译中的错误在所难免,希望读者能够多多批评指正。

让我们共同为中国的科研评价工作和文献计量学研究贡献自己的一份力量。

原版前言

本书的阅读对象为学术界,以及从事科研评价和政策研究的人员。具体来讲,本书直接面对的是如下4类读者:

一是所有曾经从事或者将要从事利用引文分析对科研绩效进行定量化评估工作的科学家和学者;

二是希望了解引文分析的基本特点、潜在功能和限制的研究政策制订者和管理人员;

三是考虑将引文分析作为评估工具的同行评议委员会成员和其他评价者;

四是定量科技研究、信息计量学、图书馆和信息科学领域的从业人员和学生。

引文分析利用引文数据构造并应用了一系列反映学术成果的“影响力”或“质量”的指标,例如学术研究论文中脚注或参考文献中的引文数据。这些指标被应用于学术交流研究和科研绩效评价。“学术”这一术语包括了科学和学术的各个领域,不仅包括通常意义上的科学,例如自然和生命科学、数学和技术科学,也包括社会科学和人文科学。

本书中的“研究政策”的意义比较宽泛,包括多层次的政策:国家负责学术研究的政府部门的科学政策、研究组织或机构进行的与质量控制和研究经费有关的研究政策、研究小组或者部门的主管开展的研究管理,包括学者的聘用、晋升等。

本书讨论了绩效研究中的关键问题:学术成果对学术知识进步的贡献。本书的主要问题是:如何恰如其分地把引文分析作为在对这种贡献的评估过程中的一种工具?尽管分析的主要部分与基础科学有关——基础科学也是引文分析应用最为频繁的领域,但是本书还讨论了引文分析在应用和技术科学、社会科学和人文科学中的应用和限制。

本文提供了广泛而重要的“事实”,纠正了一些有关引文分析的常见误解。本书介绍了基本概念和差异,既有理论层面又有技术层面,能应用于各种情况,在个人学者、研究小组、部门、机构、国家学术体系、学科或领域,以及期刊等层次都有涉及。本书揭示了对学术文献进行定量化、文献计量学分析的巨大作用,有助于更深层次地理解学术活动和绩效,并突出其政策关联性。但是,本书也有批评意见,指出了引文分析在科研评价中的限制,并对可能存在的滥用提出了警告。本书提出了将引文分析作为一种科研评价工具的正确使用的标准。

本书主要介绍了如何利用来自美国科学信息研究所(ISI)出版的“科学引文索引(SCI)”的数据及有关的引文索引(尽管美国科学信息研究所最近改名为“汤姆森科技信息集团”(Thomson Scientific),本书还是使用了其原名)。本书集中讨论了在学术交流体系研究中,特别是科研评价中使用ISI的引文索引,而不是介绍其在学术文献检索方面的突出功能。但是,本书讨论的很多方面与一般意义上的引文分析有关,与被分析的文献数据来自于哪个数据库无关。

科学和技术的定量化研究是一个正在快速发展的领域,其发展与全球学术体系中的一些

整体趋势关系密切。国家政府和研究机构需要系统化的评估,以优化研究资源分配、重新定向研究支持方向、改革研究组织、重新构造特殊领域的研究或者提高研究产出。评价文献计量学是科学和技术量化研究的一个子领域,其目的是通过对学术文献进行量化分析构造科研绩效的指标。引文分析是其关键方法之一。

在美国,ISI(即现在的“汤姆森科技信息集团”)和CHI研究^①等公司对评价性文献计量方法论做出了主要的贡献,而在其他国家,特别是欧洲和澳大利亚,这些方法在学术研究部门得到了进一步发展。本书的目的是,将评价性文献计量学当做是一种有自身的方法论和不同理论观点的多学科科学—学术活动,从而促进评价性文献计量学的学术基础。

将引文分析正确地应用于科研评价之中有较高的能力要求,并不是任何一个能使用ISI引文索引的人都能够轻易完成的。相反,本书表明,正确的引文分析需要详细的技术知识和理论理解。然而,本书并非提供如何进行引文分析的详细技术手册。相反,本书将重点放在了引文分析的主要框架、基本原则和假设,以及引文分析在各学术领域的应用和限制上。

学术团体的成员和决策者而非本书的作者决定是否使用引文分析进行评价。为了正确地做出这样的决定,关键是所有的参与人员都了解引文分析的本质、引文分析的指标是如何构建和计算的、有关衡量对象的各种理论观点,以及引文分析在学术科研评价中的优缺点。本书的目的就是提供这样一种知识。

本书的结构

本书呈现的是作者从事的部分研究,其中有一些是与荷兰莱顿大学(Leiden University)科学技术研究中心(以下简称“科技研究中心”)的同事合作完成的,并借鉴了许多领域内的其他学者的先期成果。下表概括了本书的总体结构,并指出了各部分的主要目标读者。

本书框架	主要目标读者
概述	决策者、管理人员和评价人员
第1部分:总体介绍和主要结论	学术界各领域感兴趣的学者
第2部分:理论与实践	定量科学研究从业人员和学生

第1部分包括了内容介绍,阐述了在科研评价中使用引文分析的基本观点和主要结论。另外,第1部分还包括了后面各章节的大纲,简单介绍了各章的主题和结论,但不包括大部分的技术细节。第1部分的目的是向读者,特别是非定量科学研究专业的读者介绍本书的主要内容和结论。

第2部分详细讨论了24个实践和理论章节的结论,总共分为8个单元:

单元	标题
2.1	基础科学研究部门和科技期刊的评价
2.2	ISI引文索引

^① CHI研究公司(CHI Research)是美国著名的从事专利分析研究的知识产权咨询公司。译者注。

续表

单元	标题
2.3	评价社会科学和人文科学
2.4	准确性问题
2.5	理论问题
2.6	引文分析与同行评议
2.7	宏观研究
2.8	新的发展

第 2.1 单元对后续章节进行了进一步的介绍,讨论了与引文分析类型有关的各种问题,而有关引文分析的类型在后文会被频繁使用其中包括:对基础科学研究部门的历史绩效进行评价、使用期刊影响因子和相关引文指标对期刊影响进行评价。

致 谢

我感谢以下同事对各章的初稿提出了宝贵的意见和建议:荷兰阿姆斯特丹大学的 Maarten Coolen,第 2 章;美国汤姆森科技信息集团的亨利·斯莫尔(Henry Small),第 6、7、12、13、14 章;荷兰皇家艺术与科学院的 Paul Wouters,第 15、16 章;挪威研究和高等教育研究所的 Liv Langfeldt,第 18 章;比利时佛兰德大区部(Ministry of Flemish Community)的 Marc Luwel,第 19、20 章;荷兰莱顿大学的 Jan Reedijk,第 19、21 章;美国国家科学基金会的 Rolf Lehming,第 22 章;美国芝加哥大学的 Abraham Bookstein,第 24 章。

另外,我还要感谢期刊文献的合著者,因为这些文献在本书中发挥了重要的作用:荷兰莱顿大学科技研究中心的 Anthony van Raan,第 4 章;科技研究中心的 Thed van Leeuwen,第 4、5 章;原科技研究中心、现荷兰乌得勒支中央博物馆的 Renger de Bruin,第 4 章;莱顿大学的 Jan Reedijk,第 5 章;比利时鲁汶 Steunpunt O&O Statistieken 的 Wolfgang Glänzel,第 5 章;科技研究中心的 Martijn Visser,第 8、10 章;比利时佛兰德大区部的 Marc Luwel,第 4、11 章;科技研究中心的 Anton Nederhof,第 11 章。

我还要感谢我所在的机构(莱顿大学)和 Anthony van Raan(科技研究中心主任)给我提供了撰写这本书并使用科技研究中心的资源的机会;感谢 Suze van der Luijt(科技研究中心)在本书准备阶段提供的帮助;感谢 Peter Negenborn 和 Erik van Wijk(科技研究中心)在科技研究中心建立并维护 ISI 引文索引的文献计量数据库——本书很多章节都使用了此数据库。

本书中有一重要部分是基于我 2003 年和 2004 年在科技研究中心讲授“科学评测”研究生课程的系列讲义。我向佛兰德大区部、荷兰科学研究组织(NWO)和科学出版集团 Elsevier 致谢,感谢它们对本书所提到的各部分工作开展的资助。

最后,我要感谢 Eugene Garfield(美国科学信息研究所的荣誉董事长)在本书的准备过程中给予的大力支持,本书第 12 章还包括了他对准确性问题的注释,同时他还对本书的结构和各章提出了宝贵的建议。

目 录

译者序	1
原版前言	3
概述	1
第1部分 总体介绍和主要结论	5
第1章 总体介绍	7
第2章 基本观点和总体结论	15
第3章 大纲	21
第2部分 理论与实践	45
第2.1单元 基础科学研究部门和科技期刊的评价	47
第4章 基础科学研究部门的引文分析	48
第5章 科学期刊的引文分析	61
第2.2单元 ISI 引文索引	71
第6章 基本原理、引用关系和术语	72
第7章 ISI 对各学科的覆盖面	78
第8章 科研评价中应用 ISI 引文索引的意义	92
第2.3单元 评价社会科学和人文科学	97
第9章 自然科学、社会科学和人文科学的差别	98
第10章 扩展引文分析——经济学的案例研究	102
第11章 法学科研绩效的案例研究	105
第2.4单元 准确性问题	111
第12章 准确性问题	112
第13章 引文统计的准确性	115
第14章 作者姓名、机构和子领域划分的问题	121
第2.5单元 理论问题	127
第15章 参考文献和引用能衡量什么	128
第16章 引文理论探讨——一些基础性观点	140
第17章 引文分析在科研评价中的应用意义	148
第2.6单元 引文分析与同行评议	151

第 18 章	同行评议与引文分析的运用及其有效性	152
第 19 章	对研究类院系同行评议的分析	158
第 20 章	对国家研究理事会的分析	163
第 2.7 单元	宏观研究	171
第 21 章	全球科技论文产出率在 20 世纪八九十年代增长了吗?	172
第 22 章	测度国家论文产出的发展趋势	179
第 23 章	国际科学合作是否有益	189
第 24 章	美国的科学家是否过度引用本国的论文	193
第 2.8 单元	新的发展	201
第 25 章	新指标的发展	202
第 26 章	电子出版、新型数据库和搜索引擎.....	209
第 27 章	未来的研究	213
参考文献		216

概述

本书是一本关于科学研究的书,或者更一般地来讲,是关于学术研究的书。本书的重点是那些被表征为“基础”、“基本”或“战略”的那一类研究。本书意识到这类研究对全球经济进步和社会福利的重要性,但同时也承认,要保持这类研究的稳定政治或社会基础,只有进一步建立一个内部质量控制和绩效增强的系统。本书的目的在于指出引文分析就是这个系统中的一个有用的工具。

本书主要介绍了如何评价学者通过研究成果对正确的学术知识做出的贡献。本书讨论了对单个学者、研究小组、部门和机构、学术期刊和国家学术体系的科研绩效的评价,同时分析了全球科学和学术的一般特征。

本书探讨了引文分析的作用和限制。引文分析包括反映学术成果的“影响”或“质量”的各类“文献”指标的构造和应用,它依赖于引文数据,例如学术研究论文脚注或参考文献的引文数据。本书主要研究美国科学信息研究所(ISI,即现在的“汤姆森科技信息集团”)的引文索引,但是主要的结论也适用于其他引文索引数据库。

本书的目的是,通过系统讨论学者和决策者有关引文分析价值的多种观点、纠正并指明引文分析的长处和短处(特别是相对于同行评议),为学术界的成员和制定研究政策的官员提供有关引文分析的基本技术方面的有用信息,包括引文分析衡量什么、如何将其合适地应用于科研评价和政策制定过程当中。

如果引文分析能做到以下几点的话,利用引文分析对个人、小组和机构进行评价可能更合适:

- 正式——例如在进行引文分析之前告知评价者或决策者,以及接受评价的学者或机构各种指标是评价信息的来源之一。

- 公开——接受文献计量分析者有机会检查支撑数据的准确性,并提供他们认为与得出合适的定量结果有关的背景信息。

- 学术性——文献计量学分析人员在一个学术框架下发布结果,讨论可靠性问题,明确指出理论假设,强调结果的潜在意义和限制。

- 补充有关接受评价成果实质性内容、被评价学者的工作情况和研究目的等专业和背景知识。

- 在一个明确的政策背景下实施——例如应用在一个所有参与者都清楚评价过程的角度和目标的框架下。

- 鼓励用户清楚地表明其有关学术质量、领域范围以及如何实施和加权的基本观点。

- 启发式而不是公式化的——指标被用于获得在评价过程中某一特定方面的见解,而不

是作为公式的输入、通过数学计算得到评价过程的结果。

应用引文分析评价基础科学和科学期刊以往的科研绩效已经达到比较复杂的程度。本书讨论了接受引文分析的科学家、期刊编辑和决策者提出的多个问题,并介绍这些问题能够原则上通过技术手段解决。

基于 ISI 引文索引的引文分析在多大程度上能够应用于包括应用和技术科学、社会科学和人文科学在内所有学术领域是人们经常争论的一个问题。本书仔细检查了各种学术领域的“书面交流系统的结构”之间的差异,探讨了这些系统在多大程度上被 ISI 的引文索引所覆盖。

ISI 引文索引并没有声称覆盖了全部的期刊,只是声明包括了最重要的期刊。ISI 引文索引的创始人尤金·加菲尔德提出了一种将扩大由在某个领域具有明显重要性的核心期刊组成的数据库的有力、独特的标准:期刊被索引中已经包括的期刊引用的频次。

应用一条“数据库内部”标准,本书表明,ISI 的覆盖范围在物理学、化学、细胞生物学和生物化学、与人类有关的生命科学和临床医学领域非常出色;在应用和工程科学、与动植物有关的生命科学、地理科学、数学、心理学和与其他医药和卫生有关的社会科学领域较好,但称不上出色;在其他领域包括社会学、政治学、人类学和教育学,特别是人文科学领域则一般。

部分非出色的覆盖范围的主要原因是国际期刊之外的其他来源的重要性,例如专著和会议论文。相对于科学和学术的其他领域,在 ISI 一般覆盖范围内的领域,语言或国家障碍发挥了更大的作用。另外,研究活动可能被分解为具有自身“范式”的不同的学术派别。

本书区分并介绍了 4 类文献计量学研究,其中 ISI 数据库发挥了不同的作用。在一个学科应用哪种类型的研究取决于其被 ISI 引文索引覆盖的程度。相对于具有出色覆盖范围的领域中的“标准”分析而言,在具有较好、但称不上出色覆盖范围领域的这一数据库可能通过几种方法加以扩展,而当领域的覆盖为一般时,文献计量分析可能就只能发挥有限作用,或者根本不能发挥作用。

如果研究成果在多大程度上超过了单纯国家性或地区性观点、并接受来自广泛国际学术受众的意见被认为是评价社会科学和人文科学研究质量的一个相应标准的话,一项主要的工作将是为各个子学科制定反映这方面科研绩效的有效指标。本书认为,不能想当然地认为 ISI 引文索引为这些学术领域的所有子学科提供了这些指标。一大挑战是系统化探讨其他数据来源和方法论。活跃在各个子学科的学者的专业知识和观点将在这种探索中发挥重要作用。

关于期刊影响因子,本书从技术和历史的角度解释了 ISI 影响因子是如何计算的,强调了影响它们的准确性和应用的若干问题。本书还介绍了其他一些期刊的影响指标如何解决这些问题,但同时强调不存在单一的衡量期刊绩效的“完美”指标。尽管研究小组发表文章的期刊的状态能够一部分地反映科研绩效,但是期刊影响因子不应该被用作研究小组所发表文章的引用影响的替代指标。

数据准确性是另一个关键问题。本书介绍了没有情报依据的数据采集和分析将如何严重影响引文分析的结果。使用不准确的数据不仅可能歪曲某些小组的结果,还可能影响整个文献计量研究的可信度甚至可用度。但是,准确性问题可以在处理被评价学者及其机构的过程中通过先进的数据分析手段和数据校验程序来予以解决。

下一个关键问题与“引文测量什么”有关。对基础科学研究小组的引文分析的结果大多在

统计上与对小组过去的绩效进行的同行评议成正相关关系。本书介绍的一些经验案例将表明这种正相关关系。这些结论为在科研评价中应用引文分析提供了进一步的理论依据,但是相关性也并不是完美的。

有人认为引文数能够被认为是智力影响的表现,但是引用影响与智力影响的概念并不一致。对于智力影响的概念可能存在不同的观点,而评价学术成果的评价者可能对哪些才是应该考虑的关键方面存在不同的见解。引文分析的结果必须在一个定性的、评价性的框架下进行衡量,将接受评价的成果的实质内容考虑进来。

对引用影响的理解涉及到探索可能的偏差。因此,在哪一个集总层次上进行引文分析非常关键。对各种对象的不同集总的评价可能在某种程度上忽视了单个对象的特殊特征和情况所带来的效果。必须强调,将集总看作整体的系统化偏差仍然可能发生,在研究当中应该考虑到这一点。

相对于政策环境,在同行评议的环境下利用文献计量学指标对单个学者、研究小组或部门进行合适评价的条件更容易得到满足。因此可以认为,一般来讲,在这些较低集总层次上进行的文献计量分析要达到政策支持层面最好通过同行评议。但这并不是说引文分析只是同行所使用的一个工具。

本书介绍了如何将引文分析作为一个评价同行评议程序、并保持同行评议诚信的工具。从后一个角度看,它也是决策者的工具。本书表明,引文分析有其自身的长处和短处,同行评议也是如此。问题是如何将这两种方法合适、有效地结合起来。

在“科研评价实践”研究中,一个小型同行评议委员会对整个全国性领域的研究部门进行了评价;这项有关科研评价实践的研究提出了这些评价实证是否能够确定真正优秀或“顶级”研究部门的问题。这一研究强调了研究政策制订者需要彻底反省进行这类实践的目标,同时将其成本效率考虑进来。

本研究还有证据表明,同行评级系统(例如按“优秀”、“良好”、“一般”、“较差”分)大多得到一个依赖于评级系统本身的各部门的评级分布,而这一分布在某种程度上与接受评价部门的整体绩效水平没有关系。

一项有关国家研究理事会的资助程序的研究表明,申请人与评价资助申请的专家委员的临近关系使得结果不公平。它表明了定量的、文献计量学方法能够有效帮助资助机构解决关于资助过程和评价标准的内部争论,同时也有助于解决资助机构和国家研究政策界之间的公开争论。

引文分析是致力于解决有关学术系统一般性问题的政策研究的最有价值的工具,这些问题的复杂性超过了专家委员会的能力范围。全球学术体系研究和国家学术体系的“宏观”研究就是很好的例子。本书介绍了的研究有四个方面,它们探讨了定量科学研究领域的“经典问题”,具有很高的政策相关性。这4类研究是:

- 20世纪80年代到90年代期间,科学家的全球文献产出率有增加吗?
- 怎样衡量国家文献产出的趋势?
- 国际科学合作有益吗?
- 美国科学家是否更多地引用了他们自己国家的文献?

本书介绍的第1个宏观研究讨论了20世纪80年代到90年代全球文献的产出率,其定义

为某年度单位的活跃科学家发表的论文总数。结果表明,尽管一个“普通”的科学家能够理直气壮地称自己近年来发表的研究论文比往年更多,但是从全球角度看,在过去的 20 年里科学文献的产出率并没有增加。一种解释是,对基础科学而言,同时提高科学系统的内部产出率及其经济相关性和合作对基础科学而言是自相矛盾的政策。

现在,很多国家都出版《国家科学指标报告》,分析文献计量学宏观指标如何反应一个国家研究系统的状况以及国家的科研绩效水平。然而,不同的指标和方法可能带来不同的结论,这种情况并不稀奇。这使得文献计量学指标收到选择性使用和操作的威胁。本书介绍的第 2 个宏观研究提供了如何构造和应用基于文献的宏观指标的技术性信息。

本书在评估一个国家的文献产出时试图将文献作者分为国内(在国内研究机构工作的作者)和国外(活跃在其他国家的作者)两类。所考虑的指标回答了如下问题:国家的科学从业者规模是扩大了还是缩小了?本国参与单位作者的论文数量是增加了还是减少了?本书的结论是:关键在于计算一系列的指标、并对其采取一致性的理解。将单个指标与其他指标隔离开来会歪曲结果,导致有偏差的结论。

第 3 个宏观研究讨论“国际合作有益吗”这一“经典问题”。研究结论是,当科学先进的国家相互合作时,在 10 个合作案例中约有 7 个会受益,也就是说,相对于“单纯国内性”文章,合作双方都会提高合作文章的引用影响。但是,当科学先进的国家与科学不那么先进的国家进行双边合作时——同时带动了科学在比科研评价通常观点更长的时间内的进步——这种活动可能会影响它们的短期引用影响,特别是当它们的角色位于第二位的时候更是如此。

据称美国作者过多地引用了其他美国同行的论文,这将使 ISI 引文索引挑选期刊时对美国产生偏好,将歪曲引文分析的结果。本书认为关键问题在于评价美国科学家的引用行为的标准是否合适。第 4 个宏观研究没有发现证据能得出科学领域的美国科学家过多地引用了他们自己国家的文献这一结论。

最后,本书讨论了指标发展和学术出版的最新趋势。本书强调,有必要对引文分析应用于科研评价的实际条件,以及引文分析对学术界、评价者和政策领域产生的影响进行系统化研究。这些研究成果可能促进引文分析作为一种科研评价工具能进一步发挥“关键”的潜力。

对论文和引文实践发生的变化进行分析能够给我们带来启发,但是主要问题并非学者的行为是否受到应用文献计量学指标的影响而发生改变,而是将这些措施作为科研评价工具进行运用是否提高了科研绩效,并推动学术的总体进步。

由于越来越多的学术文献能够通过互联网以电子形式获得,在不久的将来,将它们作为引文分析的来源的做法会越来越多。从科研评价的角度看,将更多的来源包括进来并不一定导致对学者促进学术知识进步的贡献的评价变得更加有效。来源文献包含新知识,以及满足专业质量标准的程度是一个关键问题。

引文分析的结果通常以一种排名的形式呈现给“外界”,例如个人学者、研究部门或机构的排名。同行评议的结果也是如此。本书认为,决策者和公众了解各研究小组学术质量的诉求是合法的,但是学术质量并不像其他很多社会领域的绩效那样能够进行直接测量和排名。不仅如此,排名忽视了某个对象的绩效与其他对象绩效之间的依赖关系。文献计量学研究学者应该在得出的结论中找到解释这几点的方法。

第 1 部分 总体介绍和 主要结论

