

〔苏〕 И. В. Маршак ова 著

# 科学引文索引与 科学前沿预测

赵云龙等 译



兵器工业出版社

# 科学引文索引与科学前沿 预 测

〔苏〕И·В·Маршакoвa 著

赵云龙 姜洪礼 曹 哲 译

曹 哲 校

兵器工业出版社

---

(京)新登字049号

## 内 容 简 介

本书精辟地论述了《科学引文索引》的基本思想、主要内容、检索功能及方法, 举例阐明了利用科学引文索引进行文献定量研究, 获得关于出版物和学者群体知识的理论和方法, 尤其详尽地论述了利用引文数据库绘制被誉为划时代成就的科学图, 跟踪科学发展, 预测各学科的科研前沿, 为科学管理者提供决策依据, 为科学研究者指示主攻方向。书中还引用了E·加菲尔德博士提供的从未公开发表过的许多资料。

本书可供科技情报工作者、图书馆工作者、各学科科研人员及管理者阅读, 也可作为图书情报专业教学的参考书。

## 科学引文索引与科学前沿预测

〔苏〕И·В·Маршакoвa 著

赵云龙等 译

曹 哲 校

科学出版社 出版发行

(北京市海淀区车道沟10号)

各地新华书店经销

中国人民解放军农牧大学印刷厂印刷

开本: 787×1092 1/32 印张: 10.5 字数: 225千字

1992年7月第1版 1992年7月第1次印刷

印数: 2500 定价: 12.50元

ISBN 7-80038-666-X/G·70

## 前 言

管理总是离不开情报检索、加工和评价。近10年来,各科学团体和行政部门都在重新审查情报资源管理的传统方法,审查科学研究和长期规划的协调机制。现代化技术的广泛采用,首先是通用计算机、个人专用计算机和电视通信的采用,以及西方主要国家所奉行的情报政策,迫使我们必须用新的观点阐明情报研究的任务,务必将此项研究与探明科学发展趋势紧密地结合起来。

情报用于科学管理,不外乎两个方面:实用方面与战略方面。从实用说,情报之利用直接关系到文献检索、事实检索以及将情报传递给研究者和行政领导人等各方面问题。人们过去和现在都是从这个目的出发来研究各种情报加工系统和检索系统(情报检索系统、自动设计系统、自动化控制系统等)。实际上这种实用观点一直束缚着情报研究,仅在新的开拓和研究中利用现有情报。

至于说到战略方面,它首先涉及的不仅是科学结构问题,或者说各知识领域的结构问题,同时还有跟踪科学的发展问题。跟踪时,最重要的是要专注于科学主题谱系的持续性变化,也就是说要把科学发展中的趋势揭示出来,而不仅是有关问题的信息。这个方面的重要意义,直到60~70年代才被少数从事科学学研究的专家所认识。

为了解决科学管理的战略任务,研究科学长期发展的趋

势是十分重要而有益的。借助文献计量方法，也就是使用参考书目情报的定量分析法，这种趋势便能揭示出来，因为参考书目情报都记录在传统的或机读的载体上，而且构成现代情报数据的基础。情报利用的战略观点直接联系到大量情报文档的算法处理，特别是联系到这些文档的群体划分，以明确科学的结构和进一步跟踪其变化。随着科学引文索引的问世，这样的研究不仅在英、美、法、日、匈牙利等国，而且在苏联也广泛地进行了。但是，由于缺乏必要的技术基础和科学引文索引系统的磁带，使得我们国家战略情报研究至今未能达到具体实施阶段。

80年代初，当第一批科学图和科学图册根据科学引文索引数据库的数据绘制成功并与大家见面之时，许多科学研究者和行政领导人开始把视线集中到情报使用的战略方面上来，而不是简单地盯住查找新情报。许多国家都大刀阔斧地开展科学引文索引系统数据库的文献计量研究，旨在评估当代科学发展水平，并为祖国科研管理、长远规划以及制定科学政策而获得必要的资料。

近几年来，在开展情报学的基础研究方面出现了某种危机。在西方发达国家里，例如英、美等国，随着情报技术成就（各种各样通用的、小型和微型计算机）的出现，却在情报学本身研究方面出现了空白。有例为证：1982年末，美国有人批评建立跨学科研究所的方案，说这个所只从情报学基础研究等等。在同一时间内，英国有人批评现代技术发展纲要，指出该纲要中没有情报学基础研究部分。西方一名研究者B·布鲁克斯在论述情报学基础研究问题时提出一个论题，说包括情报学在内所有社会科学，都需要知识量化，因

为所有科学都在飞跃前进，都在从文字说明过渡到量度论证。根据这位学者的意见，情报方面目前至少有三个主要问题必须深入研究：1) 计算机中知识的表达问题；2) 计算机处理知识问题；3) 人机关系问题。这里最重要的是使新用户熟悉并会使用计算机，实际上是用户教育问题。

第一个问题关系到情报系列化和情报浓缩。第二个问题被认为最难解决，直接关系到知识的计算机处理和筛选，即如何使计算机能自动地筛去陈旧和多余知识及事实。文献计量分析法与B·布鲁克斯提出的前两个问题有着紧密的联系。本书将在科学引文索引系统数据库基础上，来研究和论述文献计量分析法问题。

科学引文索引系统同许多别的情报系统有所不同，它除去常见的出版物书目信息（如著者、篇名、书刊名称、出版地以及著者通信地址等）之外，还把出版物的全部引文（通常称为“参考文献目录”，也有的称为“文后书目”、“引文目录”、“参考文献”、“参考书目”等等）输入数据库。关于出版物中引文的地位和意义，过去争论过，今天仍在争论。可是承认引文现象是属于科学中最主要的道德标准这一点，大家都不怀疑。

应该指出，出版物中的引文，除去道德功能之外，还应看成是科学交流的工具。留心考察科学论文的引文，能使我们及时跟踪一个科学领域的发展，跟踪它对邻近领域的渗透。D·普雷斯指出，引文能够“把全部科学论文编织成一个统一的网系。每一篇论文都是在其它论文基础之上产生的，而其本身又是后来论文的一个起点。指明出处，是一名学者继承能力的鲜明表现。从许多科学杂志中看得出来，公

开引用前辈著作的传统似乎始自1850年。从传统而言，论文是问题的深化和补充，对自己而言论文又是构思的新成果”。

论文中如果不指明真实引用来源，那就是一种剽窃行为。大家知道，卡尔·马克思是最反对任何形式剽窃的。科学引文索引系统的组织者 E·加菲尔德博士认为，引用是一种奖赏制度，是学者付给同行的硬币工资。论文中列出的参考书目，形成论文独特的上下文关系，给出文章中将要研究的问题的第一个概念，它是理解论文立言宗旨的关键。E·加菲尔德写道：“引文是科学概念的符号。这一论断构成了引文索引的理论基础”。

到70年代末，引文便成了许多知识领域的备受欢迎的研究对象，例如心理学、物理学、生物学、生物医学、社会学，而特别是情报学和科学学都重视研究引文。引文目录被当作检索工具而与传统检索工具并驾齐驱，并且又当作一种尺度，用来衡量论文的科学水平、研究者的生产力和期刊的社会效益。引文利用的实例可以举出很多：在情报领域，可以用于自然分类法，用于编制索引和文献检索。在科学学方面，包括科学计量工作在内，可以用于文献的评价和选择，科学研究能力的揭示，还可以用来分析科学学科的发展。

引文的利用可以分为三个方面来研究：1) 文献检索(实用方面)；2) 利用引文数据评估出版物或学者群体的水平；3) 利用引文从整体上确定知识或科学的结构，以便跟踪科学的发展。科学引文索引系统的建立，首先是为了满足第一项任务的需要。但不要忘记，这个拥有一大批数据库

的系统决不是一个简单的“引文统计员”，而是一部罕见的“书目参考书”，其中检索程序绝非一般，需要我们对系统的细节和规定予以充分的重视和认识。本书第一章将详细讨论科学引文索引系统的情报功能。

到目前为止，最大的争论是利用引文统计数字来评估学者的潜力问题，也就是上述第二个方面的问题。这里需要强调一下，出版物引用的数量数据，只是说明研究的成果对科学团体的影响，以及给其他学者带来什么利益。数据本身并不对出版物的质量盖棺论定。它只是一种指示器，告诉人们某著作具有较高的概率等级，因而表现出较大的意义。引文的统计数字扩大了评估出版物的可能性并使评估做得更客观一些。美国情报研究所近年来的许多统计研究确信，引文统计所发生的变化不仅是业务上的，而且也是科学现状发生变化的反映，利用统计数据来进行科学计量、评估时应想到各个知识领域论文引用都有自己的尺度和标准。许多反对采用引用数据的意见正是源于利用科学引文索引系统未得其法。

随着科学引文索引系统的出现，引文分析便有了实用价值，因为从此便有可能引用一些崭新的建立在科学引文索引系统统计数字基础之上的特征数据来描述科学的发展，这种特征数据就其规模和准确性而言都是空前的。目前科学工作者面前呈现的一项任务就是把这些新型的特征数据变成一种独出心裁的指示器，用以指明各知识领域发展的现状。D·普雷斯的研究指明，目前存在着彼此互相部分重叠的两个引文子集：引用新老“档案”文献的引文和引用“现役”的、应时的和处于科研前沿的文献的引文。只是由于以前的出版物有相当一部分与崭露头角的科学论文紧密相连，故D·普

雷斯写道：“可以把这个相当的部分也看成是活跃的研究前沿”。因此，活跃的科研前沿的概念，D·普雷斯在60年代中叶便提出来了。

目前引文的数据正在成功地用于绘制科学图（第三个方面），用于此目的的各种共引量度单位是文献被分群为数据组的一个组成部分。用来确定出版物之间联姻的共引思想是一种有效的跟踪工具，从历史角度跟踪科研前沿的发生和发展。科学引文索引系统就是实现这一思想的理想试验场。

在利用引文方面，尽管存在着一些限制条件，例如在引用时，学者面前遇到的社会心理障碍和交际障碍（首先是语言障碍），但引文分析在80年代便成为文献计量研究中的重要工具之一，它使我们有可能揭示科学发展的主要方面，非单维地揭示科学的结构，揭示各个领域的侧面，而国家级的科研课题常常会从这些侧面冒出来。如此文献计量研究的成果目前正从美国情报研究所出版的科学图和图册中不断涌现出来。由于这些新的情报资料每年都有出版，所以正如P·C·吉利列夫斯基所指出的那样：“定时跟踪科学动态，跟踪学派和科学集团的动向，因而做到有目的地影响其发展、控制其运行的时代已经到来了。在此关键时刻，加菲尔德公司的新技术和新方法又开辟了一条新路，朝着科学发展更深层次的机理积极进军”。

本书目的就是：1）研究美国情报研究所出版的主要检索刊物、机读数据库及其联机检索（第一章）；2）说明科学引文索引系统是跟踪科学发展的工具。这项任务由第二、三、四章共同完成。其中第二章主要是分析科学文献计量研究，提到情报流量化的两种方法：第一种方法用于研究科学

各个客体(出版物、著者、杂志等)的动态特征;第二种方法是用来揭示客体之间的关系,它们的分类和相互作用。这一章里还对收入科学引文索引系统中的苏联科学杂志进行了对比分析。第三章讨论了根据引用网系分析而绘制科学图的问题,详细地叙述了共引方法。这是文档算法分类的一种重要方法,它能把科学的自然主题结构全面揭示出来。第四章研究了1983~1984年的科学图,介绍了科学图册及其1990年以前的展望,指出在世界科学前进道路上各国科学贡献对比分析的可能性。

最后,本人衷心感谢 E·加菲尔德博士寄来了科学引文索引系统的直观资料并应允公布于众。这些资料我不止一次地在讲义中使用过。拙著实际上就是在讲义基础上成书的。本人还要感谢 P·C·吉利列夫斯基,没有他的帮助,本书是不可能问世的。还要感谢 B·П·莱丽格拉玛,在写作时她帮助我翻译了许多生物医学术语,并提供了许多技术性的帮助。

凡是对本书的内容提供过意见、建议和希望的朋友们,请接受本人感激之情。

# 目 录

|    |       |
|----|-------|
| 前言 | ( i ) |
|----|-------|

|              |       |
|--------------|-------|
| 第一章 科学引文索引系统 | ( 1 ) |
|--------------|-------|

|                                  |       |
|----------------------------------|-------|
| 第一节 科学引文索引系统的创建史<br>及其特点         | ( 1 ) |
| 第二节 科学引文索引系统的数据库                 | (12)  |
| 第三节 科学引文索引                       | (23)  |
| 第四节 社会科学引文索引                     | (28)  |
| 第五节 艺术和人文科学引文索引                  | (33)  |
| 第六节 杂志引文报道                       | (36)  |
| 第七节 科学评论索引                       | (51)  |
| 第八节 科技会议文献索引及社会科学和<br>人文科学会议文献索引 | (58)  |
| 第九节 目录快报和选择性情报服务                 | (75)  |
| 第十节 科学情报研究所在有机化学方面的<br>服务方式和出版物  | (92)  |
| 第十一节 科学情报研究所的机读数据库和<br>联机检索数据库   | (97)  |

|                |       |
|----------------|-------|
| 第二章 科学文献计量研究方向 | (109) |
|----------------|-------|

|                |       |
|----------------|-------|
| 第一节 情报流量化的两个途径 | (109) |
|----------------|-------|

|             |                                      |              |
|-------------|--------------------------------------|--------------|
| 第二节         | 科学文献计量研究第一种方法 .....                  | (116)        |
| 第三节         | 科学文献计量研究第二种方法 .....                  | (141)        |
| 第四节         | 杂志引文报道在文献计量<br>研究中的应用 .....          | (151)        |
| 第五节         | 科学社会学文献计量研究 .....                    | (166)        |
| <b>第三章</b>  | <b>科学图的绘制 .....</b>                  | <b>(184)</b> |
| 第一节         | 引用网系 .....                           | (184)        |
| 第二节         | 引用网系分析方法——自然分类法 ...                  | (186)        |
| 第三节         | 参考书目组合法 .....                        | (191)        |
| 第四节         | 预期联姻法 (共引法) .....                    | (196)        |
| 第五节         | 激光情报文档的群体划分 .....                    | (203)        |
| 第六节         | 科学引文索引文档的共引分群法 .....                 | (219)        |
| 第七节         | 科学引文索引系统文献分群<br>法的改善 .....           | (231)        |
| <b>第四章</b>  | <b>美国科学情报研究所出版的科学图<br/>和图册 .....</b> | <b>(247)</b> |
| 第一节         | 1983~1984年科学情报<br>研究所的科学图 .....      | (247)        |
| 第二节         | 科学情报研究所的科学图册 .....                   | (272)        |
| 第三节         | 生物工艺和分子遗传学领域<br>科学图册的文献计量分析 .....    | (302)        |
| <b>参考文献</b> | .....                                | <b>(317)</b> |
| <b>译后记</b>  | .....                                | <b>(318)</b> |

# 第一章 科学引文索引系统

## 第一节 科学引文索引系统的 创建史及其特点

科学引文索引(Science Citation Index, 简称SCI)系统由美国科学情报研究所(Institute for Scientific Information, 简称 ISI) 创办于1963年。创刊人和办刊思想的指导者 E·加菲尔德博士是 ISI 的所长。该系统在苏联有不同的名称:《加菲尔德索引》,《文献引文索引》,《科学引文索引》。

SCI 是怎样产生的呢?它是仿效《联邦谢巴尔德法令引文索引》(《谢巴尔德索引》)建立的。谢氏索引在美国从1873年开始定期出刊。英、美两国的法律界在很大程度上常以前例为立法基础,就是说,法院的判决总是要参照他们国内在诉讼历史上有影响的案例。故美国的法学家从1873年开始便使用谢氏索引。该索引果然减轻了寻找与诉讼案件有关的法律情报的劳动。谢氏索引实际上就是一个诉讼案件清单,每案都附有说明,指出该案曾引用过哪个权威性诉讼案原件,在诉讼案原件中当时所作出的判决有哪些已经修改,哪些已经废除,哪些已经变更。

1949年, A·H·塞达尔应美国专利局的需要,建议为专

利制定一个类似的索引，即给每份专利设立一个登记卡，卡片上注明该专利在以后被哪些专利引用过。1955年，《谢巴尔德索引》的副主编 W·C·阿德尔在《美国文献学》杂志上发表了一篇文章，提出了谢氏索引的原理也可以用于别的领域，当然也包括科学文献领域。该文原是按 E·加菲尔德的建议写成的，他当时正是《美国文献学》杂志的副总编。就在这一年，他也发表了题为《科学引文索引》的论文。他在该文中把后来要创建的 SCI 的结构和特点大体上都勾画出来了。

第一期 SCI 索引于1961年问世，其数据内容包括了1959～1960年出版的著作。由于它包括了生物医学和技术在内的自然知识整个领域，所以称为 SCI 系统。SCI 系统1963年开始定期出版，1964年便出现在苏联科技情报所了。国立列宁图书馆从1966年开始长期订购 SCI 的全部出版物。

ISI 出版的 SCI 系统的主题谱系逐年扩大。从1973年起，SCI 系统除自然科学、技术、农业、生物医学等领域之外，又开始广泛搜集社会科学出版物，从1978年起，又吸收了艺术和人文科学的文献。

目前 SCI 系统在上世界上享有颇高的声誉，它搜集的主要出版物来自很多国家：有资本主义国家，社会主义国家，也有发展中国家。目前 ISI 除了在费城以推销自己产品为主要服务方式之外，还有 8 个所遍布于世界各大国和地区：加拿大、西德、英国、印度、日本、台湾省和新加坡，拉丁美洲各国的服务工作则由美国加利福尼亚州洛杉矶的分所负责。

在苏联，SCI 系统（或者其单个数据库）已有20多个使

用单位，遍布全国各大城市，如莫斯科、列宁格勒、新西伯利亚、海参崴等。

SCI 系统中除去现期出版物的书目情报及其著者信息之外，还有这些出版物的引文。如前所述，引文是科学的交流工具，它可以用来跟踪科学的发展方向，提供科研课题情报，指明某文的孕育文献。既然引文的作用是向公众介绍论文之间的关系，那么 SCI 系统当然也就可以用来多途径检索自然科学、工程技术和社会科学的文献，用来描绘科学的动态结构及其发展。

著名科学史家 Z·贝尔纳指出，对他来说，索引的价值是明显不过的，因为，当他的笔头触及科学史各个方面时，他打算要做的工作 SCI 索引都做了。只不过是时序相反而已。科学出版物组成了一个相互引用的有效网系，其中的引文能够跟踪到任一事物的始源。SCI 索引把一篇具体论文同全部发表过的论文之间的关系都揭示出来，不是拘泥于一个选题或一个狭窄的知识领域，而是整个科学的任何一个门类。

SCI 系统所提供的科学情报，差不多占世界每年出版物的90%。该系统数据库每年收纳图书著录款目所占有的文献源甚富：世界主要杂志 6 800余种，书与论文集 2 000余种。在该系统中，文献代码与文献类型是统一的（例如用 D 表示 Discussion 等），详见表 1。SCI 索引中的出版物包括 100多个学科的知识领域：物理、化学、数学、医学生物科学、土壤学、农业、社会科学、心理学、法学以及政治、法律、人文科学、技术、工艺、应用科学和其它许多门类学科。

表 1 SCI系统中文献类型代码\*

| 代 码 全 称           | 代码 | 文 献 源             |
|-------------------|----|-------------------|
| Correction        | C  | 校正、增补、勘误等         |
| Discussion        | D  | 讨论、会议通信、座谈会、圆桌会议等 |
| Editorial         | E  | 社论、编辑部文章、编者按语等    |
| Biographical Item | I  | 纪念论文、悼念文章、传记等     |
| Letter            | L  | 书信、通信报道等          |
| Meeting Abstract  | M  | 演讲摘要、会议记录摘要、座谈会等  |
| Note              | N  | 技术札记              |
| Reviews           | R  | 综述、评论、书刊简介等       |
| Chronology        | K  | 年表（按时序排列的大事表）     |

\* 凡是没有带代码的文献就是普通的论文、总结或报告

在 SCI 系统中，凡是新增的文献，都要根据其引文确定其类属，登记时与先登记的同类文献放在一起。此外，凡是收录的文献，都要经过同步处理并补充其所引用过的全部文献的信息。所以 SCI 系统又是一个自组织系统。用 SCI 索引检索情报时，在时间上可以往前推进而不受所选择的出发点的限制，检索较早或较晚的文章，回溯或轮流检索都由你自便。大家知道，许多科学发明都具有边缘学科的性质，在囊括全世界出版物的 SCI 系统中，各科学门类之间没有人为的界线，而这一点恰恰是该系统的优点。SCI 索引原是在传统载体和机读载体基础上建立起来的，所以随着该系统的出现，也就出现了一种跟踪科学发展的崭新方法。这是一种具有客观定量的综合研究方法，既可用于评估各科知识领域的现状和发展，又可借以从整体上绘制出科学图和图册。

应该指出，在拉丁语范围内，SCI 系统不存在语言障碍，因为它的主要索引是用著者的姓编制的，如果出版物不是拉

丁语，例如是俄语、日语或其它语种（表 2），就将著者姓名音译为拉丁语。从表 2 看得出来，SCI 系统中的文献，用 62~65 个国家的 30 种语言出版，因此，凡是收入该系统的文献，其题目一律译成英语，而在题目之后标上语言代码，以表示该文原件的出版语种；凡是无语言代码者一律视为英语。

收录文献更详细的著录见本章第三、四、五节。

如果从情报检索系统的角度来研究 SCI 索引，则其主要优点是：囊括的出版物广，检索迅速而方便，情报分布上的跨学科性质，语言工具的简便（用自然语言检索），书目数据可靠。ISI 用书本和期刊的形式出版各种传统的（纸印的）

表 2 SCI 系统中文献的语种代码\*

| 代码 | 语 种   | 代码 | 语 种        |
|----|-------|----|------------|
| AR | 阿拉伯语  | JA | 日语         |
| BE | 孟加拉语  | MA | 马来亚语       |
| BU | 保加利亚语 | NO | 挪威语        |
| CH | 汉语    | PE | 波斯语        |
| CZ | 捷克语   | PO | 波兰语        |
| DA | 丹麦语   | PT | 葡萄牙语       |
| DU | 荷兰语   | RM | 罗马尼亚语      |
| FI | 芬兰语   | RS | 俄语         |
| FR | 法语    | SC | 塞尔维亚-克罗地亚语 |
| GA | 高卢语   | SK | 斯洛伐克语      |
| GE | 德语    | SL | 斯洛维涅语      |
| GR | 希腊语   | SP | 西班牙语       |
| HE | 希伯来语  | SW | 瑞典语        |
| HU | 匈牙利语  | UK | 乌克兰语       |
| IT | 意大利语  | XX | 多语种出版物     |

● 没有代码的文献为英语文献