

微观科技评价信息服务中个性化指标的探讨

谢华玲, 徐慧芳, 郑 菲, 陈朝晖

中国科学院国家科学图书馆, 北京, 100190

[摘要] 介绍科技评价信息服务中常用数据库及计量学指标, 主要涉及国际三大检索评价系统 SCI、EI、CPCI 和中国科学引文数据库 CSCD, 以及引入个性化绩效评价工具 ESI、H-index 和具有学科特色的评论数据库 F1000 和 MR。采用上述指标, 针对科研机构 and 科学家个人设计了相应的检索评价指标体系, 为微观科技评价信息服务引入新的学术评价指标提供参考。

[关键词] 微观 科技评价 信息服务 个性化 指标

[中图分类号] G311

Study on Individual Indexes of Microscopic Science and Technical Evaluation and Information service

XIE Hua-ling, XU Hui-fang, ZHENG Fei, CHEN Zhao-hui

(National Science Library, Chinese Academy of Sciences, Beijing 100190, China)

[Abstract] The paper introduced the commonly used databases and indexes in microscopic science & technical evaluation and information service, includes of SCI, EI, CPCI, ScienceChina; and also introduced individual tools, such as ESI, H-index, F1000 and MR. Using above indexes designed the corresponding retrieval evaluation case, in order to provide the new academic evaluating indexes for microscopic science & technical evaluation.

[Keywords] Microscopic; S&T Evaluation; Information Service; Individual; Index

在科学技术领域, 对科研机构学术水平进行定量评价的指标大致分为三类, 其一是科研产出力, 指一个科研单位的出版物数量, 包括期刊、会议论文, 专利文献和专著等; 其二是影响力、质量或效应, 指各出版物被期刊论文引用情况, 可分为总被引、篇均被引、自引和他引频次等; 其三是刊物影响因子。三类指标主要由 SCI、CPCI-S、EI、CSCD、ESI、DII 和 JCR 等数据库提供。目前, 国内现行的大学排行榜、科研机构竞争力排行榜等宏观、中观评估产品多采用上述三类指标中的三项或两项而开展^[1], 其评价模式、所用指标等并不适合微观评价的需求。微观评价产品是指在诸如院士、长江学者、百人计划、杰出青年基金等项目申报中, 或在国家重点实验室评估、职称评定、科研立项以及成果评审与报奖等特殊科研活动中, 为科研人员和科技管理人员提供的随需而动的、个性化的或可定制化的服务产品。

笔者所在单位“中国科学院国家科学图书馆查新检索中心”, 一直以来从信息服务的角度, 为广大科研用户提供了用于各类申报项目的、不同层次的、数以万计的科学研究产出力及影响力微观分析报告, 主要采用指标及数据库如表 1 所示。

表 1 科研评价中的常用指标体系及对应数据库资源

申报项目	评价指标及相应数据库						
	论著数量				被引用情况		影响因子
	SCI	EI	CPCI	CSCD	SCI	CSCD	期刊 JCR
杰出青年基金	√	√		√	√		
百人计划	√	√		√	√		
国家自然科学基金	Δ	Δ	Δ		√	Δ	Δ
国家重点实验室评估	Δ	Δ	Δ		√		

院士申报	Δ	Δ	Δ	Δ	√	√	
职称评审	√	√	√	√	Δ	Δ	Δ
年度科研统计	√	√	√				

注：上表中“√”表示必选评价指标及对应数据库，“Δ”表示可选。

在我中心历年来完成的报告中，绝大多数是基于上述传统指标完成的，但也不乏个性化的服务案例，提供了除论文数量、被期刊论文引用情况和影响因子以外的其它评价指标，如ESI排名、个人H-Index指数等，进一步体现了成果水平或评价的全面性，得到了广大科研用户的认同和赞许。本文亦对这些特殊案例进行了归类分析，以为个性化微观服务产品的设计提供参考。

1 对科研机构或科研人员的科研产出进行定量评估（ESI）

无论是科研人员、科研管理人员、学生或是图书馆员，均可采用定量分析研究绩效工具 Essential Science Indicators 对本人或所在机构的科研产出进行客观评估。如，某机构某学科在世界中的排名，所在机构发表文献的被引趋势是上升的还是下降，本领域里有哪些研究成为当前最热门的话题等等。

1.1 对科研机构的科研产出进行评估

访问 <http://esi.isiknowledge.com/home.cgi>，点击 Institutions 按钮，进入“机构评估”界面。以中国科学院为例，在检索框输入机构名称“CHINESE ACAD SCI”，点击 SEARCH 按钮，结果如图 1 所示。

Rank	Field	Papers	Citations	Citations Per Paper
1	CHEMISTRY	32,594	278,886	8.56
2	PHYSICS	25,711	166,882	6.49
3	MATERIALS SCIENCE	13,281	87,206	6.57
4	GEOSCIENCES	9,378	65,669	7.00
5	BIOLOGY & BIOCHEMISTRY	5,383	41,041	7.63
6	PLANT & ANIMAL SCIENCE	7,688	40,349	5.25
7	ENVIRONMENTAL ECOLOGY	5,190	33,091	6.38
8	ENGINEERING	6,848	30,652	4.41
9	MOLECULAR BIOLOGY & GENETICS	1,987	30,545	15.37
10	SPACE SCIENCE	3,838	25,034	6.52
11	CLINICAL MEDICINE	1,377	12,147	8.82
12	NEUROSCIENCE & BEHAVIOR	1,075	10,648	9.92
13	PHARMACOLOGY & TOXICOLOGY	1,211	9,345	7.72
14	MATHEMATICS	2,403	9,093	3.78
15	MICROBIOLOGY	1,228	8,687	7.07
16	AGRICULTURAL SCIENCES	1,340	6,868	5.13
17	COMPUTER SCIENCE	2,075	4,278	2.06
18	MULTIDISCIPLINARY	655	2,173	3.32
19	SOCIAL SCIENCES GENERAL	265	1,286	4.85
	ALL FIELDS*	124,431	868,383	6.98

图 1 中国科学院进入全球科研机构前 1%的学科

由上图可以看出，按照总被引次数排序，中国科学院进入全球科研机构前 1%的学科有 19 个，分别为化学、物理、材料科学、地学等学科。由上图可知，最近 10 年间中国科学院的上述 19 个学科被 Web of Science 收录的论文总数、总被引次数和篇均被引次数，以及在各学科的高被引论文及热点论文，学科排名等信息。以地学为例，中国科学院地学学科在最近 10 年间，按照总被引次数在全球科研机构中排名第 4。

1.2 对科研人员的科研产出进行评估

访问 <http://esi.isiknowledge.com/home.cgi>，点击 Highly Cited Papers (last 10 years) 按钮，进入“高被引论文”界面。以著者 ZHANG PZ 为例，在检索框输入科学家名称“ZHANG PZ”，点击 SEARCH 按钮，结果如表 2 所示。

表 2 ZHANG PZ 的论文在 GEOSCIENCE 学科中的引用频次和排名情况。

出处	学科领域	引用频次	排名
SCIENCE 294 (5542): 574-577 OCT 19 2001	GEOSCIENCES	252	162(of 2533)
GEOLOGY 32 (9): 809-812 SEP 2004	GEOSCIENCES	164	427(of 2533)

NATURE 410 (6831): 891-897 APR 19 2001	GEOSCIENCES	147	561(of 2533)
CHINESE J GEOPHYS-CHINESE ED 51 (4): 1066-1073 JUL 2008	GEOSCIENCES	48	1761(of 2533)
SCIENCE 322 (5903): 940-942 NOV 7 2008	GEOSCIENCES	18	2272(of 2533)

注：引用频次和排名情况每两个月更新一次。

“Highly Cited Papers (last 10 years)”反映的是在最近10年里发表的论文。将10年来的文章按学科、按引用次数由高到低排序，排在前1%的论文，就是高被引论文。它可以用来了解具有长期影响的一些研究。由上表可以看，ZHANG PZ有5篇论文在GEOSCIENCE领域进入“近十年来高引频论文”。

2 获取特定研究人员的引文报告和H指数(WOS)

Web of Science(简称WOS)不仅可以检索科学研究的由来与发展,通过数据库的“Create Citation Report”功能,还可以生成引文报告(Citation Report)和查询H指数(H-index),这使Web of Science成为了解研究业绩的重要工具之一。H-index指数是美国加利福尼亚大学圣地亚哥分校统计物理学家乔治·赫什(Jorge E. Hirsch)于2005年发明的旨在衡量科学家个人/团体科研成就的一种评价指标^[2],其巧妙地将数量指标(发文数)和质量指标(被引频次)结合在一起,克服了以往各种评价科学工作者科研成果的单项指标的缺点。一个人的h指数是指他至多有h篇论文分别被引用了至少h次, h指数越高,表明他的论文的影响力越大。例,查询机构为“University of Science and Technology of China”的作者HOU JG的引文报告和H指数^[3]。

访问<http://apps.isiknowledge.com>,进入Web of Science数据库,在检索框中输入Author=HOU JG and Address=UNIV SCI TEHC CHINA,点击SEARCH按钮。

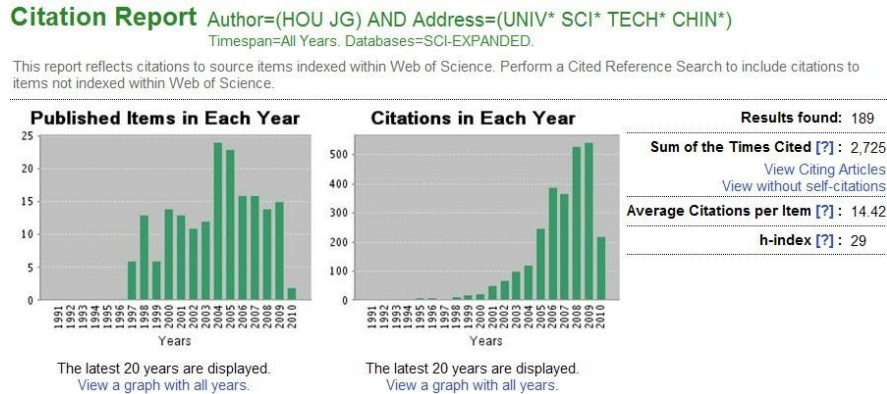


图2 作者被收录论文、论文被引用情况年代分布图示和H指数

由上图可以直观获取论文收录、被引用情况和H指数。另外,通过观察上图中每篇论文的年被引情况,用户还可以了解目前哪些研究成果正被成为关注重点。赫希认为如果一个人从事科研20年后(从发表第一篇论文算起),h指数能达到20,就算是一名成功的科学家,而HOU JG的H-index为29。

3 查询论著被会议论文库收录论文引用情况(CPCI)

2008年10月,ISI公司将其国际会议论文数据库ISTP和ISSHP分别更名为CPCI-S(Conference Proceedings Citation Index-Science)和CPCI-SSH(Conference Proceedings Citation Index-Social Science & Humanities),并对它们新增了原来只有SCI和SSCI才具有的引文检索功能(其可检索出论著被期刊论文引用情况),用来检索和分析论著被CPCI收录的会议论文引用情况,如表3所示,对特定学科,如计算机、数据库

技术，选择 CPCI 作为引用源更具优势。

表 3 计算机技术、工程数学类论著被 SCI、CPCI-S 引用情况对比

出处	发文数	SCI 引用	CPCI 引用
中科院计算所 LI XW	28	36	50
中科院数学所 CHENG DZ	152	758	405

4 学科评论类数据库

学科类数据库，是相对于综合类数据库而言，具有学科领域特色的资源数据库，其一般具有检索文章、获取评论信息或查询引文等功能，如生物医药领域评论库(Faculty of 1000, F1000)、美国数学评论(Mathematical Reviews, MR)、空间科学领域引文库(Astrophysics Data System, ADS)和高能物理学领域引文库(SLAC's SPIRES-HEP, SPIRES)等。本文仅以评论类数据库 F1000 和 MR 为例，引入到微观评价服务中，主要是获取其针对论文给出的独特评论信息。

4.1 生物医学类评论数据库 F1000

F1000 (<http://www.f1000.com/>) 包括生物学 F1000 和医学 F1000，是由生物医药领域的顶级专家将最近一个月内生物学领域的重要论文或可能改变医学实践的极少数优秀论文(包括新奇发现、技术进展、有趣假说、重要确认和争议性发现)，通过评述、评分(其中，3分-推荐、6分-必读、9分-优秀)和汇总排名后推荐给科研人员，让他们了解本专业最新的进展。专家为每篇文章写一个简短的评价，说明文章的主要贡献和入选理由，不考虑期刊的状况^[4]。如图 3 所示，该篇文章得分为 6 分，为科研人员必读类文章。

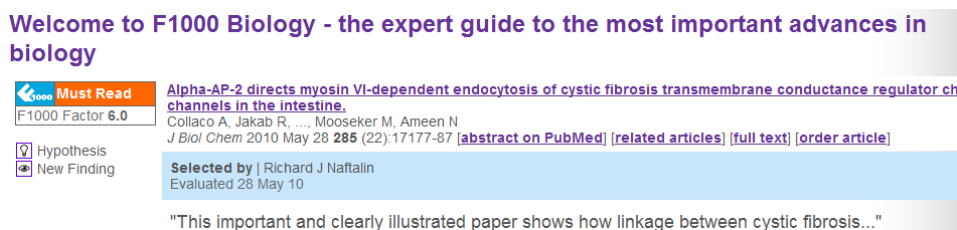


图 3 论文著录、被评信息及得分情况

4.2 数学领域评论数据库 MR

美国数学学会(American Mathematical Society, AMS)于 1940 年推出的权威数学评论数据库《Mathematical Reviews》，是数学家根据数学学科分类，委任世界各地的专业评论家对最新出版的数学和应用数学领域研究文章(包括：期刊、图书、会议录、文集和预印本)作评。其对 400 多种数学核心期刊做全评，目前，中国有 150 多种期刊被选评。如文章“Partial upwind finite element method for time-dependent convection-diffusion equations”的被评信息：In this paper, two upwind finite element methods are studied for the two-dimensional linear time-dependent convection-diffusion problem. One is……will be feasible and they will preserve the discrete maximum principle (Reviewed by Zhiyong Zhao).

学科评论类数据库打破了单纯以引文数或杂志的影响因子来评价一篇文章质量的传统方法，为评估高水平生物学、医学或数学等领域学术成果提供了新的评价指标。

5 科研机构或科学家个人竞争力分析指标设计

依据上述传统和个性化指标，设计了科研机构或科学家个人竞争力分析或学术水平的定量评价指标体系，机构可以是研究所或实验室，指标体系构建情况如表 4 所示。

表 4 竞争力分析或学术水平评价指标体系构建表

论著数量	SCI	EI	CPCI-S	CSCD	ESI	F1000	DII	专著
第一和通讯作者论文比例								
高影响因子论文比例		—	—				—	—
论著影响力 (SCI 引用)	引用次数	年均引用次数	篇均引用次数	H-index	高频引用次数比 (<10/10-30/30-50/50-100/>100)			
其他引用	SCOPUS	CSCD	Google Scholar	DII	ADS	MR	Citeseerx	
学术部门影响因子(2或5年)								
论文产出效率	人均产出 (篇/人)			投入产出 (篇/百万元)				
国际合作	主要合作国家及合作网络图			主要合作机构及合作网络图				
代表性成果数量	单篇引用次数大于 100 的论文数			发表在某领域内期刊影响因子排名前 10 位的高影响因子论文数				
人员配置	学术带头人	学术带头人/其它科研人员		科研/支撑/行政		人员规模 (人)		
学科分布								
获奖情况								
科研经费								
科研设施								

需要强调的是尽管上表中涵盖了大部分文献计量学指标,但在实际运用时并非意味着要罗列表格中的所有指标,而是需要根据评价对象、评价目的以及用户的要求,从中选取具有重要意义的指标即可。

另外,分析案例不能仅从科研产出及影响力角度来建立评价指标体系,在实际运用中,还需结合科研投入产出效率、研究开发方向、人员配置、科研设施、科研经费与项目、学科领域重要奖项获得等其它经济指标或社会指标来全面分析和评价某机构或某科学家的整体状况,这样制定出来的发展策略才具有科学性、实用性和可操作性。

6 科研机构竞争力分析案例设计与分析

针对机构作具体分析时,应在充分调研目标机构自身特色等情况的基础上,从表 4 中选择相应指标进行组合分析。同时,在分析中需要寻找对比机构进行对比分析,通过分析可借鉴对比机构的核心或优势竞争力,来优化目标机构的发展战略、建立赶超目标等。

本案例在依据美国周刊等权威数据基础上,除目标单位外还分别选取了欧洲、美洲、亚洲等环工领域的数个领先研究机构以进行竞争力对比分析,机构包括 Eawag、Stanford University 环境工程专业、日立国立环境研究所、清华大学环境科学与工程学、浙江大学

环境工程系、哈尔滨工业大学市政环境工程系。选取的主要指标包括：2000-2008 年期间论文引用状况（引用总数、年均引用频次、单篇引用次数、H-index、高频次引用论文比例）；论文分布（期刊分布、影响因子分布、通讯作者比例）；论文产出效率（人均产出、每万元发表论文数量）、人员配置等。部分数据和统计结果如下所示：

表 5 2000-2008 年各机构 SCI 引文分析结果

项目	EAWAG	NIES	STANFORD	清华环工	浙大资环	哈工大	目标单位
平均引用次数/年	2402	2283	825.3	345.6	414	79.5	307.5
篇均引用次数*	15.93	9.71	12.87	6.5	5.62	3.65	7.54
引用总频次	21618	20547	7428	3110	3732	716	2768
SCI 论文总数	1745	2644	774	675	974	384	555
通讯作者比例(%)	56	42	75	75	75	72	81

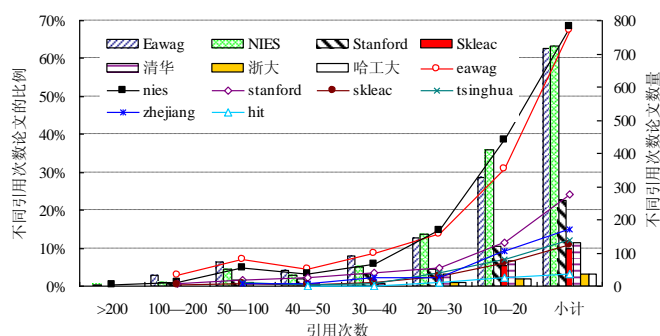


图 4 不同引用次数的论文占论文总数的比例

主要对比分析结论：在对本案例中采集的各机构各类数据进行统计和对比分析基础上，结合目标单位实际情况，为该单位提出的发展战略目标如下：

名称	论文单篇平均引用次数(次/篇)	人均论文产出(篇/人)	引用次数>10 以上的论文占论文总数的比例(%)	IF>3 期刊论文占论文总数的比例(%)	人均经费(万元/人)
指标	10-12	1.8	30	35	100

名称	学术带头人(人)	学术带头人/其它科研人员的比值	科研/支撑/行政(%)	人员规模(人)
指标	19	1:5~1:7	85/10/5	112~156

7 结语

上文对传统的和目前一些比较热门的数据库或评价指标进行了简单的归类，并给出了部分服务产品的设计思路和案例。虽然这些资源在国内外文文献中已有较多研究或应用报道，但将其引入到个性化的用户服务方面的报道甚少。在评价学术成就时，无论传统的同行评议制度，还是系列计量学指标，作为一种单一的情报学工具和评判标准，均具有一定程度的局限性。实际运用时需要两者或更多的指标进行综合计算与分析，才能得到逼近客观的、可供参考的数据。

参考文献：

- [1] 蒋国华. 科研评价与指标. 北京, 红旗出版社, 2000
- [2] Hirsch, JE. An index to quantify an individual's scientific research output. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, 2005, 102(46):16569-16572

- [3] 美国汤姆森科技信息集团(Thomson Reuters)中文页面之“怎样获得某个研究人员的引文报告和H指数?” [2010-05-25] <http://www.thomsonscientific.com.cn/scholarlyresearch>
- [4] 陈美伦. 一种新的医学论文学术评价系统——医学F1000. 中华医学图书情报杂志. 2008, 17(2):70-72

作者简介: 谢华玲, 女, 1982年生, 馆员, 发表论文2篇; 徐慧芳, 女, 1977年生, 馆员, 发表论文近十篇; 郑菲, 女, 1971年生, 副研究馆员, 发表论文10余篇; 陈朝晖, 男, 1969年生, 研究馆员, 发表论文10余篇。