

国际合作论文中机构贡献度分析

——以中国科学院为例

刘筱敏 崔剑颖 何莉娜

中国科学院国家科学图书馆 北京 100190

〔摘要〕 利用2006-2010年SCI数据,以中国科学院国际合作论文为对象,分析中国科学院5年中国际合作论文的状况,通过国际合作论文中机构的数量,计算合作出版指数,对中国科学院在国际合作论文的贡献度进行量化计算,进一步分析中国科学院在国际合作中的参与程度。

〔关键词〕 国际合作 机构贡献度 出版指数 文献计量学 期刊论文

〔分类号〕 G350

Institutions Publishing Index of International Collaboration

——Taking Chinese Academy of Sciences as An Exmaple

Liu Xiaomin Cui Jianying He Lina

National Science Library, Chinese Academy of Sciences, Beijing 100190

〔Abstract〕 International cooperation plays an important role in scientific research. Papers as the results of international cooperation research can show how an organization participates in international cooperation. In this paper, 2006-2010 SCI data are used to review the status of international cooperation research papers. The analysis focuses on the contribution of Chinese Academy of Sciences to international cooperation papers. The number of international cooperation organizations is used to calculate the co-published index. The index quantitatively analyzes the degree of Chinese Academy of Sciences' contribution to international cooperation papers so that the degree of Chinese Academy of Sciences' participation in the international cooperation is analyzed further.

〔Keywords〕 international collaboration contribution degree publishing index bibliometrics journal papers

当前,国际合作渗透到国家发展的各个层面,科学研究之间的国际合作亦是国际合作的重要内容之一,国际合作论文的产出折射出一个国家在国际合作中的地位和参与程度。2011年3月英国皇家学会出版了一份题为 *Knowledge, Networks and Nations: Global Scientific Collaboration in the 21st Century* (以下简称 KNN) 的报告,报告中提及,在国际期刊上发表的论文中有35%以上是国际合作的产物,这个数字在15年里增加了10%,其中以含有美国科学家在内的国际合作的增长率为最高——以美国科学家为主要作者的合作论文的发表数量从1996年的约5万篇上升到2008年的约9.5万篇,但美国的国际合作论文数量仅占到其2008年论文总量的29%。与之相比,德国2008年论文总量有几乎一半都是来自国际合作研究。中国的国际合作研究也在增加,但国际合作论文数量占其论文总数的比例,从1996年的18%下降到2008年的15%^[1]。KNN的报告描述了国际合作在科研活动中不断增加

的现象,说明了国际合作对科学研究活动的推动作用。

各国国际合作论文的增加说明科研领域中国际合作强度的增加。论文数量是分析各国在科研领域参与国际合作程度大小的特征值之一,而一个国家作者在国际合作中的贡献度更能说明其在国际合作的地位。传统的文献计量学统计一篇论文不同作者的贡献度,是通过赋予不同作者的权重值来计算的,比较有代表性的有三种:第一种方法为划一计数法:将一篇论文的名誉全部归于第一作者,这种方法是目前最常用的也是简单易操作的方法;第二种方法为正规计数法:不论作者在一篇论文中的顺序如何,每个作者都被计算为产出一篇论文,近几年这种方法经常用于第一作者与通信作者两者论文贡献度的计算;第三种方法为调节计数法:根据一篇论文作者的数量进行等额分配,如果一篇论文有两个作者,则每个作者的分值为1/2,如果一篇论文有三个作者,则每个作者的分值为1/3,依次

类推^[2]这种方法在某个机构的科研绩效评估中经常得到应用。

英国纳皮尔大学计算学院的 Elisabeth Davenport 和美国印第安纳大学的 Blaise Cronin 在 *Journal for American Society for Information Science and Technology* 2001 年第 9 期上发表论文“Who Dunit? Metatags and Hyperauthorship”, 该论文通过分析 ISI 数据库中合作论文不断增长的趋势, 对作者论文的贡献度进行了讨论, 提出了不同层次深度对作者贡献度的计量方法, 计算方法渗透在创作、引用、检索、传播的整个流程之中, 全面考察了在论文生命周期中每个作者对论文的贡献度。但受到每个作者的文字责任清晰划分、引用论文的章节或段落明确定位等各种条件的制约, 目前仅从论文载体形式和内容的表现形式方面, 暂时还无法对每个作者的贡献状态进行准确的描述, 因此无法得到一个完满的结论。

为了计算一个国家或一个机构在论文中的贡献度, *Nature* 于 2008 年推出了一种出版指数的计算方法, 计算同一机构的作者数量占该论文作者总数的百分比^[3], 将其百分比赋予每个作者所属机构作为该机构对该论文的贡献度的分值。根据机构的参与人数计算某个机构对论文的贡献程度, 更为客观。在一定程度上避免了单纯平均分配每个机构贡献的计算方法, 同时也避免了论文的成果只统计第一机构而忽视其他机构的贡献的现象。

我国的学者对国际合作论文的数量及特点也多有研究。通过 SCI 论文统计, 发现在基础研究领域, 我国参与国际科技合作的科研主体主要集中在高校和专业科研院所。中国科学院以其庞大的科研规模和实力, 成为我国专业科研院所参与国际科技合作的典型代表^[4]。

本文以中国科学院国际合作论文为分析对象, 从论文数量到作者贡献度两个方面表述中国科学院在国际合作中的地位与特点。

1 数据来源及统计方法

本文使用的数据采集于 2006 - 2010 年 SCI-CD (SCI 光盘版), 以“Peoples-R-China”为检索词, 下载我国发表的论文。所分析的论文包括 article、letter、review 三种类型。将数据下载后, 对每篇论文的机构进行处理, 以含有中国科学院直属研究院所以及中国科学技术大学的论文集合作为统计对象。中国科学院直

属近百个研究所, 论文作者的机构区分粒度以研究(院)所为最小单元。

2 中国科学院国际合作论文概况

中国科学院是我国重要的科学研究机构之一, 历年的数据统计说明, 中国科学院直属研究院所及中国科学技术大学在我国科学研究国际传播中占有重要的地位, 在历年 SCI 数据统计中, 论文产出量名列前茅^[5]。

2006 - 2010 年, 中国科学院以全部作者统计, 发表论文 77 864 篇, 其国际合作论文 21 751 篇, 占全部论文数量的 27.93%。这个比例高于 KNN 的报告。该报告中提到 2008 年中国全部国际合作论文占全部论文产出的 15%, 在相同的统计窗口(2008 年), 笔者统计中国科学院的国际合作论文比例为 26.98%。两个比例值不同之处在于 KNN 的数据样本来自于 Scopus, 统计数据的样本量与本文的统计样本量不尽相同, 分析 Scopus 的期刊国家分布特点得知, 在国际重要期刊的收录上, Scopus 与 SCI-CD 有一致性, 但 Scopus 收录我国出版期刊的数量(543 种)远远高于 SCI-CD 收录的我国期刊数量(20 种), 根据中国科学引文数据库(CSCD)数据统计, 国际合作论文在我国期刊上刊登的比例很低, 以 Scopus 为统计依据, 由于论文总数增大, 导致我国国际合作论文比例会低于 SCI-CD 数据统计的比例, 因此两者不具有可比性。KNN 中提到 2008 年美国国际合作论文的比例为 29%, 考虑到美国发表论文的期刊具有多元化及国际性特点, 且中国科学院论文多刊登在国际期刊上, 虽然统计的样本量不同, 但是这个数字仍然具有一定的可比性, 中国科学院国际合作论文比例接近美国国际合作论文的比例。

2006 - 2010 年间, 中国科学院发表国际合作论文数量呈年度增长趋势(见表 1), 论文数量增长幅度最大的年份在 2008 年——比 2007 年增长了 16.23%。相比而言, 第一作者国际合作论文年度增长幅度略低于全部作者的论文增长幅度——2008 年比 2007 年增长了 12.08%。

以第一作者统计, 中国科学院第一作者的国际合作论文占全部作者国际合作论文的百分比始终维持在 47% 左右, 并有年度小幅下降的现象: 2006 年第一作者的国际合作论文数占全部国际合作论文的 47.67%, 2010 年这个比例下降到 46.04%, 中国科学院研究人员作为合作者的论文数量略高于作为第一作者论文数量。

表 1 中国科学院国际论文数量年度分布

年度	全部国际合作论文数(篇)	其中第一作者国际合作论文数(篇)	第一作者/全部作者百分比(%)
2006	3 740	1 783	47.67
2007	3 759	1 854	49.32
2008	4 369	2 078	47.56
2009	4 935	2 319	46.99
2010	4 948	2 278	46.04

对中国科学院国际合作论文进行进一步分析, 2006-2010 年第一作者发表论文总计 51 977 篇, 第一作者国际合作论文总计 10 312 篇, 国际合作论文占总论文量的 19.84%。从年度百分比来看, 2006 年为 19.07%, 2010 年为 20.63%, 逐年略有上升。该比例低于全部作者的国际合作论文占全部作者论文数量的比例, 说明以中国科学院作者为主导的国际合作论文数量尚需进一步加强合作。

在以上陈述是通过中国科学院研究人员在论文署名中的不同位置说明国际合作数量的分布状态。以目前流行的方式, 将论文的成果归于署名第一的机构, 这种方法简单易行, 但忽视了合作者对论文的贡献, 在国际合作中, 每个机构都是论文成果的贡献者, 应该对非第一作者机构的贡献度予以计算。

3 合作出版指数

依据 *Nature* 出版指数的计算方法, 对中国科学院在国际合作论文中的贡献度进行计算和详细分析。

下载的 SCI 论文数据不能得到作者与机构一一对应的结果, 因此无法准确统计一篇论文中每个机构所拥有的作者数量, 为此直接使用 *Nature* 的出版指数计算方法会造成偏差。虽然在下载的论文集中无法得到每个机构的作者人数, 但可以获取一篇论文的机构个数——相关著录会详细到同一研究院的不同研究所、同一所大学的不同院系, 表现了一个机构的最小单元, 因此可以借鉴 *Nature* 出版指数的计算方法, 对论文所属机构的论文贡献度名誉进行分配, 目的在于测算中国科学院在国际合作论文中的作用。本文将该分值称为合作出版指数。合作出版指数的计算方法为: 同一机构在同一篇论文中出现的次数除以该篇论文机构总数。合作出版指数越高, 说明在论文中的贡献度越高; 反之, 说明在论文中的贡献度越低。

本文的研究中, 当中国科学院为论文第一作者机构时, 定义为第一作者论文; 中国科学院为非第一作者机构时, 定义为合作者论文。经过计算, 得到中国科学

院每一篇论文的合作出版指数: 第一作者的合作出版指数总分为 4 710.71, 合作者的合作出版指数总分为 3 563.04, 第一作者的合作出版指数的总分值是合作者的 1.32 倍。由此得到一个结论: 虽然第一作者的论文数量小于合作者的论文数量, 但合作出版指数却高于合作者论文的合作出版指数。

笔者进一步将合作出版指数划为 4 个区域, 每 0.25 划分为一个区域, 出版合作指数大于 0.5 为高合作出版指数区域, 小于 0.5 为低合作出版指数区域, 如表 2 所示:

表 2 中国科学院论文合作出版指数分值统计

CAS 合作出版指数	>=0.75	<0.75 and >=0.5	<0.5 and >=0.25	<0.25
第一作者论文(篇)	290	6 003	3 410	609
百分比	2.81	58.21	33.07	5.91
合作者论文(篇)	20	2 748	5 524	3 147
百分比	0.17	24.02	48.29	27.51

注: CAS 为中国科学院英文名称缩写形式

通过表 2 可以看出, 中国科学院作为第一作者机构发表论文时, 有约 2/3 的论文合作出版指数大于等于 0.5, 说明中科院对论文至少有 50% 的贡献度。38.98% 的论文合作出版指数小于 0.5, 说明当中国科学院作为第一作者时, 合作出版指数较高, 对论文的贡献度较好。中国科学院作为合作者时, 合作出版指数整体偏低, 3/4 的论文合作出版指数分值小于 0.5, 仅有 20 篇论文的合作出版指数大于 0.75, 占合作者论文数的 0.17%, 对论文的贡献度较低。

在上述计算中, 虽然区分了第一作者与合作者, 但没有对第一作者给予权重处理, 出现了第一作者论文中最低合作出版指数为 0.011 6 的情况, 这是因为该论文的合作机构有 86 个, 而中国科学院仅有一个机构出现, 因此所占份额较低。在合作论文中, 最低的合作出版指数为 0.004 8, 该论文的合作机构多达 208 个。如果考虑到国际合作论文的作者署名顺序, 强调第一作者的重要性, 可以对第一作者、合作者赋予不同的权重, 再计算合作者的合作出版指数, 这种方法比较适合不同机构之间的比较。

合作出版指数的计算与合作规模有很大的关系, 如果合作的规模比较大, 合作机构比较多, 合作出版指数的分值则较小。因此, 笔者对合作机构规模进行统计, 方法是统计一篇论文的合作机构数量。总体上看, 国际合作论文集中于中国科学院与一个机构的合作(见表 3), 这种现象在第一作者论文中尤为明显——

在第一作者论文中有 50% 以上的论文为与一个机构合作而产生的,合作规模较小,因而合作出版指数的分值会比较高。合作者论文的合作机构规模大于第一作者论文的合作规模,合作出版指数偏低。从第一作者论文的合作机构数量分布看,随着合作机构数的增加,合作论文数快速下降(见图 1)。国际合作机构的多少从一个角度反映了合作的广度,可以说明国际合作的规模。中国科学院作为合作者,合作论文分布虽然也较为集中,但整体分布较第一作者更为均匀,说明作为合作者论文,国际合作广度更大,间接反映了作为第一作者的国际合作还处于小规模的状态,而作为合作者则进入了较大规模的国际合作中,但还没有占据主导地位。

表 3 不同合作规模论文数量分布

合作机构数(个)	1	2	3	4	5	6	7	8	9	>=10
第一作者论文(篇)	5 139	3 002	1 259	516	202	69	27	19	11	68
合作者论文(篇)	2 309	3 383	2 296	1 301	629	345	144	86	65	881

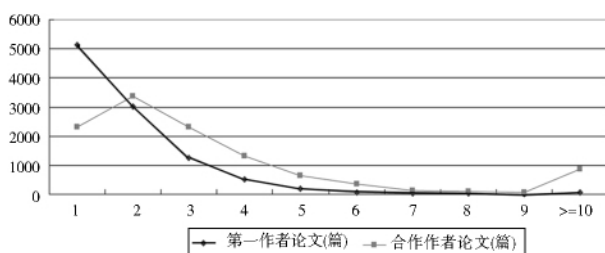


图 1 不同合作规模论文数分布

4 结 语

(作者简介) 刘筱敏,女,1965年生,研究馆员,发表论文 20 余篇;崔剑颖,女,1988年生,硕士研究生;何莉娜,女,1988年生,硕士研究生。

(上接第 53 页)

- [9] 陈成鑫,初景利. 国外新一代用户网络信息行为研究进展[J]. 图书馆论坛,2010(6):71-75.
- [10] 曹梅. 略论用户信息行为研究的演进[J]. 图书情报工作,2010(1):55-58.
- [11] 邓小咏,李晓红. 网络环境下的用户信息行为探析[J]. 情报科学,2008(12):1810-1813.

(作者简介) 茹意宏,男,1966年生,副教授,发表论文 30 余篇;彭爱东,女,1971年生,副教授,发表论文 20 余篇;黄水清,男,1964年生,教授,发表论文 50 余篇。

虽然在我国科研活动中的国际合作已经具有了一定比例,通过以上对中国科学院 SCI 国际合作论文的统计,可以看到国际合作论文的数量占 SCI 总论文数量的比例年度分布比较稳定,国际合作绝对论文量呈年度上升状态。通过计算合作出版指数的方法,测度中国科学院在国际合作论文中的贡献度,分析论文数量上升时,机构对论文的贡献程度。数据显示中国科学院为第一作者时,机构贡献度值较好,但论文的国际合作规模较小。中国科学院为合作者时,机构贡献度较低,但论文的国际合作规模较大,说明了在国际合作中,不仅要保持论文数量的年度增长,更应该关注在论文成果中的贡献度,特别是应加强在大规模的国际合作科研活动中作用,提高其论文的贡献度。

国际合作的科学研究受到学科特征、参与方式、国际合作背景、国际合作机制及投入等多方面因素的影响,因此对我国国际合作需要进行更多角度的分析,以更为客观地反映其发展状态。

参考文献:

- [1] The Royal Society. Knowledge, networks and nations: Global international collaboration in the 21st Century [EB/OL]. [2011-04-28]. <http://royalsociety.org/policy/projects/knowledgenetworks-nations/report/>.
- [2] Nature Publishing Index China [EB/OL]. [2011-09-08]. <http://www.natureasia.com/en/publishing-index/china/>.
- [3] 丁学东. 文献计量学[M]. 北京:北京大学出版社,1993.
- [4] 刘娅. 基于文献计量的我国基础研究领域国际合作态势分析[J]. 中国科技论坛,2010(3):149-154.
- [5] 中国科学引文数据库项目组. 中国科学计量指标: 论文与引文统计(2010年卷)[M]. 北京:知识产权出版社,2010.

- [12] 王艳,邓小昭. 网络用户信息行为基本问题探讨[J]. 图书情报工作,2009,53(16):35-39.
- [13] 彭文梅. “信息行为”与“信息实践”——国外信息探求理论的核心概念述评[J]. 情报资料工作,2008(5):33-36.
- [14] Nied-wiedzka B. A proposed general model of information behaviour [EB/OL]. [2011-10-25]. <http://Informationr.net/ir/9-1/paper164.html>.