

外部特征不规范对论文收录的影响及改进措施

——以某大学论文为例

张 勤, 蒋纯国

(上海理工大学 图书馆, 上海 200093)

摘要: 用文献计量学的方法, 通过 Web of Science 检索对 2006—2009 年期间 SCI 收录某大学论文的外部特征著录规范性情况进行综合统计; 对照 SCI 的论文统计规则, 实证分析投稿过程中在论文外部特征著录规范方面的不足所产生的负面影响, 以引起广大科研工作者重视; 为提高国际论文的收录及检索效率提出建议。

关键词: SCI 收录; 国际论文; 外部特征; 著录规范; SCI 统计规则; 规范化建议

中图分类号: G 252.7 **文献标志码:** A **文章编号:** 1009-895X(2011)04-0309-05

The Effects of Irregularities of External Characteristics on Paper Collection and Measures for Improvement

—Taking a University's Papers as an Example

Zhang Qin, Jiang Chunguo

(Library, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: The normality of external characteristics of university teachers' international papers recorded by SCI during 2006—2009 will be counted by Web of Science Search with bibliometrics methods. Guided by the paper statistic rules of SCI papers, an empirical study is done on the negative effect of deficient external characteristics of papers on their recording and retrieval, thereby alerting scientific researchers to this problem. We expect to improve efficiency of paper recording and retrieval.

Key words: SCI collected; international paper; external characteristics; recording standardization; SCI statistical rules; standardization advice

一、研究背景

(一) 期刊论文外部特征

文献的特征包括文献的外部特征和内部特征^[1]。外部特征是指文献上显而易见的、一般不反映文献实质意义的那些特征, 如书名(刊名、篇名)、著者和各种符号(专利号、标准号)等。内部特征是指文献

实质意义的特征, 如主题词和分类号等。期刊论文的外部特征包括: 论文的外部特征、刊登该论文期刊的外部特征。论文的外部特征有论文题名、作者、第一作者的工作单位或联系方式等; 刊登该文期刊的外部特征有期刊刊名、年、卷、期、该论文在当期期刊上的起止页码及 ISSN(国际标准期刊号)等。通常, 文献的外部特征也是检索和获取该文献的主要

途径之一,只有正确著录才有利于文献的检索与利用。

(二) SCI 简介

SCI(Science Citation Index 科学引文索引)是由原美国科学情报研究所(现汤姆森科技信息集团,简称汤姆森科技)编辑出版的世界著名的期刊检索工具,于1961年创刊,现为双月刊^[2]。它们与该研究所编辑出版的“社会科学引文索引”(Social Sciences Citation Index,简称SSCI)、“艺术与人文引文索引”(Arts & Humanities Citation Index,简称A&HCI)和“会议论文引文索引”(Conference Proceedings Citation Indexes:即原ISI Proceedings数据库,简称CPCI,分为自然科学版“CPCI-S”和人文社会科学版“CPCI-SSH”)共同组成了一个国际范围的大型综合性文献检索系统。目前,这些引文索引不仅有印刷版、光盘版和联机数据库版,而且发展了网络数据库版,可以在Web of Science上实现数据整合,提供多种在线服务。

(三) SCI 的论文统计规则

每年,中国科技信息研究所(简称中信所)会公布全国高校三大检索工具收录科技论文的统计数据,其检索规则如下^[3]。

规则 I 数据库产品限定在SCI-E,而不是全部WOS或WOK。

就SCI而言,涉及三个层面的概念:Web of Knowledge(WOK)、Web of Science(WOS)、SCI-E。其中WOK是一个检索平台,或者说是解决方案,包括很多不同的数据库和其他资源,比如WOS、Derwent、Medline、JCR、ESI等等;WOS是一个数据库集合,包含了SCI、SSCI、CPCI等等数据库产品;SCI是一个数据库产品。因为每个单位购买的数据库产品不同,所以大家在WOK平台下看到的数据库不完全一致。在检索的时候要注意不是选择所有数据库,而是仅仅限定在SCI-E这一个数据库中检索。

规则 II 目前数据库的版本是扩展版SCI-E。

SCI数据库有很多介质的版本,从最早的印刷版开始,到后来的光盘版(SCI-CDE),再到后来的扩展版或称网络版(SCI-E)。2000年以前的论文统计数据是来源于光盘版,2000年以后的统计数据则来源于扩展版。两者的区别主要是收录文献数量的差别。光盘版大约3000多种,扩展版大约8000多种。整体上看,扩展版覆盖了光盘版的大部分内容。需

要注意的是,二者实际上是分开加工的两个不同的数据库,并不是严格意义上的包含关系,并非光盘版收录了,扩展版就一定收录。

规则 III 检索时间用入库年度作为标准。

SCI-E的检索入口有两个时间的选项,一个是出版年,另一个是入库时间。出版年就是指文章发表的年份,入库时间是指论文被SCI-E数据库收录的时间。为了保证统计工作的连续性,每篇论文不会被遗漏,也不会被重复统计,SCI论文统计的时间界定采用入库时间,即收录时间。

规则 IV 论文单位要著录“中国”,统计中归属于第一作者的第一个单位。

首先,论文的作者单位在著录时一定要出现“中国”字样,否则不可能计入中国论文的统计。其次,一篇论文可能有很多作者,一个作者可能著录几个单位,论文在统计中计入第一个作者的第一个单位。有的时候论文的第一作者单位和通讯作者单位不同,出于各种原因,在SCI-E下载的数据里,往往会把通讯作者的单位提前到第一个,即论文在统计中记入通讯作者的第一个单位。

规则 V 只有四类文献计入统计:Article、Review、Letter、Editorial Material。

SCI对于期刊的选择比较严格,但被选入后的期刊上的各种类型的文献都将被收录进数据库。其中既包括学术论著这一类的文献,也包括更正、会议文献等其他信息。因此,SCI只选择内容详尽、叙述完整、著录项目齐全的文献类型进行各单位论文数的统计。这些文献类型是Article、Review、Letter和Editorial Material,其他类型的文献则不计。

二、研究意义

(一) 提高被收录的几率

美国科学情报研究所(ISI)编辑部每年都要评审近2000种新期刊,充实和调整SCI收录的期刊种数。被评估的期刊中仅10%~12%被选用,退回率达88%;此外,ISI数据库中的现有期刊也在不断筛选更新,每年尚有150~200种已入选的期刊被中途淘汰^[4]。因此,期刊论文的质量对于期刊来说显得尤为重要,每一篇论文的优缺点代表了期刊的整体水平,也是影响该期刊收录的因素之一。

近年来SCI收录的期刊数每年略有增加,如2006—2009年SCI分别收录了6166、6426、6620、7387种期刊。大约在每年的6月20日,ISI会公布

上一年度 SCI 收录的期刊名单。

作者向 SCI 源期刊投稿时应注意:查找与自己稿件的专业方向相吻合的期刊,再根据这些期刊的影响因子(IF)选择适合自己稿件水平的期刊;尽量选择用英文撰写论文;重视参考文献的引用;注重论文格式的标准化及所在单位英文译名的标准和统一,以期提高论文被收录的几率^[5]。

(二) 提高被引的几率

SCI 引用情况分析是评价基础性研究领域科研成果的一种比较客观的、定量的指标之一。被科学界同行广泛认可的原创性研究成果,其发表的论文被引次数高,这也是原创性研究特点所决定的。一方面,目前我国原创性研究成果相对较少,这就要求科研管理决策机构和基金资助部门要大力提倡和鼓励开展原创性研究工作,加大在本土培养有创新能力人才的力度。另一方面,现有的原创性成果论文要保证其在基本出版标准、编辑内容等方面的准确性、规范性,才容易被人通过各种不同途径检索得到并加以引用,从而体现该原创性成果在学术上的相关性,横向上的对应性,纵向上的继承性^[6]。

(三) 提高被统计的几率

每年中信所都会公布全国高校三大检索工具收录科技论文的统计数据并加以排名。有很多单位在科研评价中使用中信所的论文统计结果,是出于对中信所多年来论文统计结果科学性和权威性的认可。但是当所在单位公布统计结果以后,有的研究人员发现自己检索到了某篇论文,而统计结果里没有。在一些单位的科研评价体系中,论文有否被 SCI 收录与个人利益密切相关,因此出现质疑就需要安排很多时间和人力逐篇进行复检和回复。而复检的结果不是统计错误和漏检,而是由于研究人员的论文不符合 SCI 的论文统计规则,才会出现检索结果的不一致^[3]。

因此,注重国际论文格式的标准化及所在单位英文译名的规范和统一,减少错误信息,对于提高各个高校乃至整个国家的收入论文统计数,都将起到正面和积极的作用。

三、当前不规范现状与实证分析

周英智在“科技期刊作者单位英文译名存在问题及建议”中探讨了科技期刊论文中,作者单位英文译名常见的问题有:1)单位地址及名称不全;2)作

者单位译名不一致;3)字母大小写不规范。并对产生的原因进行了分析,提出过操作建议^[7]。为具体了解与分析当前教师与科研人员国际论文外部特征的著录规范遵循情况,笔者通过 Web of Science 检索,穷尽其检索策略,对某大学 2006—2009 年被 SCI 收录的论文进行了统计,并在参考该大学科技处每年公布的校内科研工作汇编资料的基础上,得到了该大学学术论文在 2006—2009 年间被 SCI 收录的实际论文数量的结果。将上述检索结果与中信所的论文统计结果逐年地进行了对比,出现了检索结果的不一致,详见表 1。

表 1 2006—2009 年 SCI 收录某大学论文数对比

Tab.1 Comparison of the number of papers for SCI included a university from 2006 to 2009 篇

收入年份	笔者检索得到的论文数量	中信所的论文统计结果数量	未被统计计入数量	未被计入论文占比 (%)
2006	54	50	4	7.40
2007	87	82	5	5.70
2008	107	95	12	11.20
2009	117	113	4	3.40
合计	365	340	25	6.80

结果显示,在 2006—2009 年期间,该大学共有 25 篇(占 6.8%)SCI 收录论文没有出现在中信所的论文统计结果中。为了查明原因,笔者分离出这 25 篇未计入文献,逐篇检索 Web of Science 分析其外部特征著录信息,发现其第一作者的第一单位著录不规范,使其不符合 SCI 的论文统计规则。分析原因如下。

(1) 第一作者的第一单位项目著录不规范,不符合 SCI 的论文统计规则 V。

25 篇未计入文献中有 11 篇文献为第一作者的第一单位名称著录不规范,即著录的某大学英文译名不规范;有 4 篇文献邮政编码信息错误;有 1 篇文献没有第一作者的第一单位信息。笔者在校对原文后发现,4 篇论文的原始文献中没有邮政编码信息,1 篇论文的原始文献中没有单位信息。因此,笔者推测:不规范的单位英文译名被 ISI 误认为其他高校,错误的邮编是因著录需要而后添加的,论文或许也被统计进了其他高校。

以上情况反映出部分国际论文著录项目的不规范、不齐全,不符合 SCI 的论文统计规则 V,故这部分论文未被统计计入。

(2) 论文在统计中进入了通讯作者的第一单位。

25篇未被统计文献中有14篇文献,它的第一作者单位和通讯作者单位不同。按照SCI的论文统计规则IV,这时候在SCI-E下载的数据里,往往会把通讯作者的单位提前到第一个,即论文在统计中进入了通讯作者的第一单位。故也未被统计计入(见表2)。

表2 25篇未被统计文献的原因分布

Tab.2 Reasons of 25 papers not in statistics 篇					
收入年份	未被统计文献数量	原因1		原因2	
		第一单位名称不规范文献数	占比(%)	进入通讯作者的第一个单位文献数	占比(%)
2006	4	2	50	2	50
2007	5	2	40	3	60
2008	12	4	33	8	67
2009	4	3	75	1	25
合计	25	11	44	14	56

为总体了解该大学的国际论文在外部特征著录方面的规范遵循情况,笔者还对2006—2009年间中信所提供的该大学被SCI统计计入的论文进行了外部特征著录分析。结果是:被统计计入的340篇论文中共有17篇文献的第一作者单位著录项目不完整(没有邮政编码)、有4篇文献出现了错误邮编信息、有3篇文献出现了作者署名格式错误。

综合以上两部分共计365篇SCI论文的统计分析,发现该大学国际论文的外部特征著录不规范问题确实存在,且具有一定比例,主要体现在第一作者单位英文译名的规范性、完整性以及著者署名的正确性等方面。在除去通讯作者计入第一单位的14篇文献外的351篇论文当中,出现不规范英文译名的论文达11篇(3%)、单位信息不完整的论文17篇(5%)、著录出现错误的论文7篇(2%)。

四、国际论文外部特征的著录规范化建议

我国加入WTO后,科技期刊在国际学术合作与学术交流中的作用日趋明显,英文版科技文献是国际交流的最好方式。因此,国际论文的编排与著录必须符合国际化、规范化、标准化的要求。针对当前我国高校国际论文中存在的问题,提出如下几点规范化建议。

(一) 作者单位英文译名要规范

作者单位英文译名在整篇论文中所占篇幅很少,

其作用也往往不被人重视,但它对科研单位的学术地位和国际论文的编排质量却有重要的影响。为了全面提高国际论文的质量,使之在国际交流中更好地发挥作用,必须做好国际论文中作者单位英文译名的规范化工作^[7]。笔者建议做好以下几方面的工作。

1. 提高对本单位英文译名规范化重要性的认识

科研论文往往是一个单位科研水平的重要反映,科研论文被国际检索工具收录也是衡量一个单位科研水平的重要标志之一。作者单位英文译名不规范容易给国外读者造成误解。为此,作者单位一定要重视英文名称在国际交流与合作中的作用,将正确使用本单位的英文名称作为一件严肃的事情来对待;单位在合并或成立新机构时,在确定中文名称的同时也应确定其英文名称;作者单位应要求其职工投稿时使用统一的单位英文译名,并做好相应的审核工作。

2. 作者撰稿时应正确使用所在单位的英文译名

有的作者在投稿时,往往仅凭自己的理解或者参考别人的用法将单位的中文名称翻译成英文,从而导致同一个单位有多种译名,甚至出现同一个作者在撰写不同的论文时也使用不同英文名称的现象。因此,每个作者都必须以严谨的工作态度,正确使用所在单位统一的英文译名,避免随意性。

3. 制定作者单位英文译名的编排格式

科技期刊要拥有更多的国外读者,编排上必须与国际标准接轨。国外的每种科技期刊在编排作者单位名称及所在地址时都有相对固定的格式。当然,编排格式也不尽一致,但总的来说,大多数期刊列出了较为详细的作者单位,有的甚至标注了所在单位的街道名称及门牌号码。因此,国内科技期刊编辑部门在编写征稿简则时应应对作者单位英文名称的格式有统一、详细的要求,让作者共知共用。

4. 编辑工作中应注意作者单位英文译名的规范化

作者单位英文译名很少涉及晦涩的语法及生僻的词句,只需稍加留意即可发现其中的问题。因此,请编辑在加工中注意作者单位英文译名是否规范,把好作者单位英文译名编排规范化这一关。

(二) 作者单位署名要完整规范

作者单位署名要完整,一般应包括:单位全称、

详细地址、邮政编码和国别等信息。特别要注意的是:1) 国际论文的单位名称中一定要出现“中国”字样,否则不可能计入中国论文的统计;2) 论文中的单位名称要规范和统一,要注意著录信息的完整性,不能缺失、更不应该有错误。

(三) 通讯作者署名的建议

由于学校与学校间、国家与国家之间的合作与交流,每年都有很多的合作科研和合作论文产生,一篇论文可能有多个作者,一个作者可能著录几个单位。通常,一篇论文只能被统计一次,如果一篇论文由多人合著,SCI 在论文统计中计入的是第一个作者的第一个单位。

通讯作者往往是整个课题总的负责人,承担课题的经费、设计、文章的审阅和把关。他也是文章和研究材料的联系人,担负着文章可靠性的责任。一般情况并不是要求所有的论文都一定要写通讯作者。通讯作者多数情况和第一作者是同一个人,只有在通讯作者与第一作者不一致的时候,或者当论文属于某课题、某重大项目的重要内容之一,而论文作者又不是课题总负责人时,就有必要标注通讯作者。因此,如果课题的立项、具体研究和撰写论文主要由第一作者完成,则第一作者应不要随便将其他人作为通讯作者^[8-9]。

出于各种原因,论文的第一作者单位和通讯作者单位有可能不同,这时候第一作者可以提供论文原件的首页(印刷版期刊上的论文的第一页,可以看出刊名、题目、第一作者和单位著录等信息)作为修改的依据,并积极向本单位科研管理部门反映情况,以取得他们的帮助。很多单位的做法是,科研管理部门每年将类似情况收集在一起,于当年 SCI 收录结果发布之前统一集中提交给中信所。该所在核实后会在最后统计中进行修正。需要注意的是信息提交的时间,例如 2010 年收录的论文只在 2011 年的 8 月份以前修正,过了这个时间,统计结果已公布,也就不再修改了。

五、结束语

综合以上统计发现,高校教师国际论文投稿过程中在论文外部特征著录规范化方面存在着不足,主要体现在作者单位英文译名的不规范。希望借本研究成果的发表,能提醒广大高校科研论文的作者,充分了解 SCI 对收录论文的外部特征、特别是著录的要求情况以及所存在的问题。这些在平时并不被高校教师所重视的因素,已经严重影响到国际知名检索工具的统计结果。因此,注重国际论文格式的标准化及所在单位英文译名的规范和统一,减少错误信息,对于提高各个高校乃至整个国家被国际有关统计机构收录论文的统计数量,将起到正面和积极的作用。

参考文献:

- [1] 滕胜娟,蓝曦主编.现代科技信息检索[M].北京:中国纺织出版社,2007:4-5.
- [2] 王细荣,韩玲,张勤.文献信息检索与论文写作[M].第2版.上海:上海交通大学出版社,2009:94-97.
- [3] 马峥.论文统计规则(SCI 篇)[EB/OL].(2010-05-24)[2011-03-15] <http://bbs.sciencenet.cn/home.php?mod=space&uid=396900&do=blog&id=327147>.
- [4] 梁红妮,胡德华.SCI 源期刊的分析研究[J].现代情报,2004,(5):166-169.
- [5] 昌兰,彭敏宁,陈丽文,等.SCI 及向其源期刊投稿的注意事项[J].中南大学学报(医学版),2005,30(5):370-372.
- [6] 周辉,张光红,蔡晖,等.原创性研究成果的 SCI 引用分析[J].中国科学基金,2002,16(2):85-87.
- [7] 周英智.科技期刊作者单位英文译名存在问题及建议[J].编辑学报,2003,15(4):265-267.
- [8] 屈轶主编.科技英语写作[M].北京:军事谊文出版社,2008:68-70.
- [9] 殷国荣,杨建一编著.医学科研方法与论文写作[M].北京:科学出版社,2009:156-164.