

两种图书馆学期刊论文关键词的统计分析

安秀芬, 林朝英, 邵胜春
(华北煤炭医学院图书馆, 河北 唐山 063000)

摘要: 以《图书情报知识》和《大学图书馆学报》为样本, 统计了2001~2003年两刊的论文关键词, 并对其数量、内容、标引质量等进行了研究分析。

关键词: 图书情报知识; 大学图书馆学报; 关键词; 统计分析

中图分类号: G254 **文献标识码:** A **文章编号:** 1002-1248 (2005) 05-0170-03

Statistical analysis on the Keywords of Papers in two Kinds of Library Journals

AN xiu-fen, LIN Chao-ying, SHAO Sheng-chun

(Library, North-China Coal Medical College, Tangshan, Hebei 063000, China)

Abstract: Based on Knowledge of Library and Information Science and Journal of Academic Libraries, this Paper gave a statistic on the keywords in this two journals 2001~2003, and studied on the number, the subject content and the indexing quality

Key words: Knowledge of Library and Information Science; Journal of Academic Libraries; keywords; statistical analysis

1 前言

哪些问题, 为两刊以后办的更好提供一点依据。

期刊论文关键词标引, 是我国文献出版发行工作标准化、规范化建设与国际接轨的需要, 也是促进我国学术研究成果走向世界的必由之路。关键词标引数量的多少, 标引质量的优劣, 直接影响着期刊质量和科研成果的检索效率^[1]。本文对2001~2003年《图书情报知识》(以下简称《知识》)和《大学图书馆学报》(以下简称《学报》)的论文关键词进行了统计分析, 旨在比较一下进入21世纪两种刊物在关键词标引数量上比1997~1999年^[2]有什么变化; 两刊在研究内容上有哪些异同点; 在关键词标引质量上还存在

2 结果分析

2.1 两刊关键词数量分布

关键词是为了文献标引工作而从报告、论文中选取出来的, 用以表示主题内容、信息、款目的单词或术语^[3], 是论文主要内容的概括。经统计, 2001~2003年《知识》共发文552篇, 标有关键词的论文为100%, 共有关键词2043频次。《学报》共发文421篇, 标有关键词的论文占总发文的91.07%, 共有关键词1340频次, 详见表1。

表1 关键词数量统计

项 目	图书情报知识				大学图书馆学报			
年 度	2001	2002	2003	合计	2001	2002	2003	合计
发文篇数	116	215	221	552	141	146	134	421
标关键词论文数	116	215	221	552	122	135	129	386
百分比 (%)	100	100	100	100	86.5	92.5	96.3	91.7
关键词频次	431	797	815	2043	418	471	451	1340
篇均关键词数	3.72	3.71	3.69	3.70	3.43	3.49	3.50	3.47

收稿日期: 2003-08-16

作者简介: 安秀芬 (1952-), 女, 副研究馆员, 发表论文25篇, 参编著作3部; 林朝英 (1958-), 男, 馆员, 发表论文13篇; 邵胜春 (1958-), 女, 馆员, 发表论文15篇。

由表 1 可看出:《知识》2002、2003 年的年发文量比 2001 年几乎增长了 1 倍,原因是自 2002 年该刊由原来的季刊改为双月刊,开本也改为大 16 开,使其载体容量大增;标引关键词的论文比例达到了 100%,比文献[2]提高了 33 个百分点;篇均关键词数为 3.70 个,比文献[2]的 3.65 个略有提高。可见编辑部对论文关键词的标引工作引起了高度重视。《学报》2001~2003 共发文 421 篇,比 1997~1999 年总发文量 528 篇减少了 107 篇,分析其原因,发现该刊随着论文质量的不断提高,论文的篇幅也在不断增加,其篇均页码由 1997~1999 年的 3.1 页增加到现在的 3.9 页;标引关键词的论文比例也呈逐年上升趋势,

年平均为 91.7%,比文献[2]的 82%上升了近 10 个百分点,但与《知识》相比还有差距;篇均关键词数为 3.47 个,比文献[2]的 2.74 个有了明显提高,但还低于《知识》的 3.7 个。

2.2 两刊不同数目关键词的论文分布

目前,学术期刊均要求作者在论文中附 3~8 个关键词。关键词提供的多少对揭示论文主题的程度往往有直接的影响。过少,对论文内容揭示不全,检索时造成漏检,会降低检全率;过多,会降低检准率。因此,关键词的标引数量隐含着一个标引深度的概念,适当的标引深度应该是在文献的检全率与检准率之间达到某种平衡。表 2 是两刊不同数目关键词的论

表 2 不同数目关键词的论文分布

关键词个数	0	2	3	4	5	6	7	8	10	合计
《知识》论文篇	0	32	245	173	74	20	4	3	1	552
数占总论文数%	0	5.80	44.38	31.34	13.41	3.63	0.73	0.54	0.18	100
《学报》论文篇数	35	41	184	117	31	10	1	2	0	421
占总论文数%	8.31	9.74	43.71	27.79	7.36	2.37	0.24	0.48	0	100

文分布状况。

由表 2 可知,《知识》标引 3 和 4 个关键词的论文最多,二者占 75.72%,比 1997~1999 年的 49.1%提高了 26.62 个百分点。《学报》由 1997~1999 年标引 2~3 个关键词的论文最多,改为标引 3~4 个关键词的论文最多,二者占 71.5%,提高了 16.66 个百分点,由此可以看出,标引 3~4 个关键词的论文无论是《知识》还是《学报》都占了论文的大多数,说明一般情况下,3~4 个关键词基本能够满足一篇论文对主题内容充分揭示的要求,符合论文关键词选取数量的一般规律。从表 3 中也看出,两刊中的极少数论文存在不足标引,《知识》的个别论文存在过度标引现象,有待进一步规范。《学报》还有 8.31%论文未标

关键词。中科协规定“没有关键词的论文应列入非学术论文类”[4]。因此,《学报》对登载的所有学术论文应都标上关键词。标关键词论文的多少和关键词个数的多少虽然不能说明期刊论文水平的高低,但可以反映期刊的规范化程度[5]。

2.3 关键词的频次分布

一个关键词出现的频次等于附有该关键词的学术论文数。因此,一个关键词出现的频次越高,相关的研究成果数越多,研究内容的集中性就越强[6]。《知识》2001~2003 年共有自然标引的关键词 1 107 个,2 043 频次;《学报》有 915 个,1 340 频次,其分布见表 3。

从表 3 可以看出《知识》关键词频次的范围是

表 3 关键词频次分布

频次	个数《知识》	个数《学报》	频次	个数《知识》	个数《学报》	频次	个数《知识》	个数《学报》
1	855	765	11	1	2	27	1	0
2	125	94	12	2	1	28	1	0
3	48	22	13	1	1	3	0	1
4	19	6	14	2	1	37	0	1
5	11	12	15	1	0	39	1	0
6	11	5	16	1	1	40	1	0
7	10	0	17	1	0	43	0	1
8	5	1	21	1	0	45	1	0
9	4	0	22	1	0	72	1	0
10	1	1	23	1	0	合计	1107	915

1~72。频次大于10(含10)的有19个关键词,占关键词总个数的1.7%,频次计451次,占总频次的22.1%,词均23.7次。这19个关键词是:信息网络(10)、网络出版(11)、信息资源(12)、高校(12)、电子商务(13)、网络(14)、图书馆管理(14)、搜索引擎(15)、企业(16)、知识管理(17)、对策(21)、建设(22)、发展(23)、数字图书馆(27)、网络环境(28)、信息服务(39)、中国(40)、高校图书馆(45)、图书馆(72)。《学报》关键词频次的范围是1~43。频次大于10(含10)的有10个关键词,占关键词总个数的1.1%,频次计199次,占总频次的14.9%,词均为19.9次。这10个关键词是:网络环境(10)、资源共享(11)、图书馆管理(11)、CALIS(12)、数据库(13)、元数据(14)、大学图书馆(16)、高校图书馆(32)、图书馆(37)、数字图书馆(43)。可以认为,这些关键词表述的研究内容是《知识》和《学报》3年来研究的核心内容,在很大程度上影响着研究的前沿动态,揭示了研究内容之间的主要内在关系,体现了《知识》和《学报》的办刊宗旨。

《知识》关键词频次少于10的有1088个词,频次计1592次,词均1.5次。其中有855个关键词频次为1,125个关键词频次为2。《学报》关键词频次少于10的有905个词,频次计1141次,词均1.3次。其中有765个关键词频次为1,94个关键词频次为2。这些关键词的频次极低,这从一个侧面反映《知识》和《学报》的研究内容具有明显的分散性。分散性的研究内容拓宽了学术研究的选题视野,有助于形成更多的研究生长点,促进研究内容甚至学科间的交叉、渗透、互容、互用,推进学术研究向更广泛的领域发展。

2.4 关键词内容分析

在统计范围内《知识》和《学报》有相同关键词225个,占《知识》关键词总个数的20.23%,占《学报》关键词总个数的24.59%,如果按关键词频次计算,《知识》是877频次,占总频次的42.6%;《学报》是546频次,占总频次的40.60%,也就是说《知识》所刊载的论文约有40%与《学报》刊载的论文主题内容相同。这些内容主要是:高校图书馆、数字图书馆、数据库、网络环境、图书馆管理、计算机编目、资源共享、信息服务、信息资源、知识管理、情报工作等。《知识》除以上内容外,还较多刊载了编辑出版发行、档案工作、教育研究、企业管理、商

业信息等方面论文,这与该刊设有相关栏目有关。由所统计关键词的内容,体现出了《知识》“集图书情报、信息管理、出版发行于一刊,融学术性、知识性、实用性、可读性于一体”的办刊宗旨。《学报》侧重报道高校图书馆业务研究中的高水平论文,研究的焦点都与大专院校图书馆事业建设中的重点、难点和热点问题相关。刊载的论文除与《知识》相同内容外,还较多报道了有关核心期刊研究、联合编目、元数据等方面的论文。由此看来,两刊既有共同的研究报道重点,又有不同的侧重。

2.5 关键词标引质量分析

关键词标引是对文献主题高度提炼和概括的创造过程。简洁、准确的关键词使论文主题突出,一目了然,起到“画龙点睛”的作用。《知识》和《学报》论文关键词的质量从整体来说较高,但仔细研究,发现还存在以下几方面问题:①两刊都存在对多名称事物的标引不规范现象。同是表示一种事物却有几种不同标引形式,如“采访”、“采访工作”;“采购”、“采集”、“图书采购”等,同样的例子很多。②《知识》通用词标引过多。所谓通用词,是指那些没有独立检索意义的泛指词,如:中国(40次)、发展(23)、建设(22)、对策(21)、应用(9)等。相比而言,《学报》标引的通用词明显较少,中国(1次)、发展(2)、建设(1)、对策(1)、应用(1)。③组配不当或不够。如“试析图书馆调研工作的内容与形式”一文(《知识》2002年2期76页),关键词标引为“图书馆、调研工作、内容、形式”,笔者认为后两个词应组配,标引为“调研内容、调研形式”更为准确。④对自由词的选择太随意。选择自由词作关键词标引时,必须是词义完整、概念准确、词形简练、实用性强、能表达论文主题概念的词汇,但在两刊关键词中,有些自由词选择不够准确,如:《知识》中的“爬行”、“前端”、“外包”等;《学报》中的“本体”、“流转”、“统治”等,这些词概念不完整,涵义不清,令人费解。

3 结语

进入21世纪,《知识》和《学报》在论文关键词的标引数量上比1997~1999年都有了较大提高,《知识》达到100%的论文标引,《学报》达到91.7%;两刊标引3~4个关键词的论文占了大多数;两刊在研究内容上既有集中性,又有分散性,既有共

(下转第176页)

算,《荟要》占《全书》的三分之一弱。基于各自繁复不同,因而编排分类也有不同。

(2)《荟要》成书在前,《全书》的分类得以在其基础上加以完善,从而更趋合理。

(3)《全书》编者从分类实践中,获得了分类经验,如不按书名分类,而按图书内容归类,按作者著书主旨归类等,这些都使得《全书》分类基本反映了我国古代的知识结构体系,以至于影响至今^[6]。

参考文献:

- [1] 四库全书荟要[M].台湾世界书局影印稿藻堂本,1988.
- [2] 刘乃和.《四库》研讨会论文集序[C].中国首届《四库全书》研讨会论文集,1994.
- [3] 欧阳修.新唐书·艺文志[M].北京:中华书局,1975.
- [4] 王兴亚.《四库全书荟要》的文献价值[J].黄河科技大学学报,2000,(3).
- [5] 丁银燕,颜作辉.四库全书总目浅探[J].图书馆,2003,(3).
- [6] 周汝英.《四库全书总目》分类法述评[J].温州师范学院学报,1999,(4).

(上接第 169 页)

2.6 网络、纸质重复报道的文献以纸质文献为主

目前,各个国家和地区、越来越多的企事业单位或部门,都在搞上网工程。随着“上网工程”的快速发展,越来越多的纸质文献上了网,借助网络这一媒体来传递、流通、阅读。笔者认为,对于那些网络和纸质重复报道的文献,在著录参考文献时还是应该以纸质文献为主要来源依据。凡是有纸质文献原文出处的,都要以纸质文献的责任者、题名和原文出处为著录参考文献的依据。例如,现在,在清华同方的中国全文期刊数据库中,人们可以查找、阅读 7 000 余种杂志上发表的近 120 万篇论文资料,在作为参考文献著录时,还是应该以原来发表论文的纸质期刊文献为主,不要著录长长的网络地址。确实无纸质文献出处

的,以网络文献出处为著录参考文献的依据。

参考文献:

- [1] 王淑华,王亨君.谈网络参考文献的著录[J].编辑学报,1999,(3):169-170.
- [2] 杨海平,等.网络参考文献著录研究[J].大学图书馆学报,2001,(1):76-78.
- [3] 刘雅轩,关家麟.网络参考文献的标引[J].情报学报,2002,(4):428-432.
- [4] 蔡宇宏.网络环境下参考文献的主要功能及其实现[J].图书情报工作,2003,(10):22-25.
- [5] 邱敦莲.农业学报参考文献著录格式调查及存在问题分析[J].农业图书情报学刊,2000,(增):17-18.

(上接第 172 页)

同的重点,又有不同的侧重;在关键词标引质量方面,总体水平较高,但还存在一些问题。针对关键词标引中存在的问题,希望期刊编辑人员在繁忙的工作中抽出一定的时间检查和修改作者标引的关键词,保证关键词的标引质量。提高了关键词的质量,也就等于提高了期刊的质量。

参考文献:

- [1] 王思哲,郝玲.关于图书情报学期刊关键词的几个问题[J].图书情报工作,2001,(6):93-95.

- [2] 刘彦庆.两种图书馆期刊关键词特征的探讨[J].情报科学,2000,18,(9):822-824.
- [3] 学位论文和学术论文的编写格式[S]. GB7713-87 科学技术报告.
- [4] 关于在学术论文中规范关键词选择的规定(试行)[J].图书馆建设,2003,(1):7.
- [5] 雷晓庆,刘晓雁.《图书情报工作》论文关键词特征的统计分析[J].图书情报工作,1998.
- [6] 安秀芬,黄晓鹏,张霞.期刊工作文献计量学学术论文的关键词分析[J].中国科技期刊研究,2002,13,(6):505-506.