

医学期刊关键词的标引

陈秋莲

(重庆市卫生信息中心 400014)

摘 要 关键词是科技论文的文献检索标识,能反映文章主题概念的词或词组,关系到文献的被检索利用情况。文章从关键词的选取范围、选取原则以及首标词的选择三个方面,探讨了如何合理选择关键词,并就关键词标引中存在的漏标、错标、复标问题进行阐述。

关键词 医学期刊;关键词;关键词标引;文献检索

中图分类号 G255.2

文献标识码 B

文章编号 1006-1525(2011)01-0056-03

Keyword indexing for medical journals

CHEN Qiu-lian

(Chongqing Sanitation Information Center, Chongqing 400014, China)

Abstract: Keyword indexing for scientific documents can reflect words or phrases indicating the theme of the document, which directly connects the information retrieval. This article expounds the principle, the scale of selection of the first key words, and how to rationally select key words, how to avoid repetitions, mistakes and omissions of key words.

Key words: medical journals; keyword; key word indexing; document retrieval

关键词是科技论文的文献检索标识,能反映文章主题概念的词或词组,包括主题词和自由词两部分。主题词是专门为情报检索机构编制索引用的,它是从自然语言的主要词汇中选取后并加以规范化的词或词组;自由词是尚未规范化的词或词组。关键词标引的数量及质量影响其在数据库中被检索利用的频次,进而关系到期刊的影响因子、总被引频次等。许多期刊规定文献的关键词要有3-8个。

1 关键词的选取范围

关键词可从文章的题名、摘要、层次标题或文章其他内容中抽取。题目中的专业词汇在很大程度上概括了主题的含意,所以,首先可考虑从篇名中选用,题目中所含的具有实指意义的信息应在关键词中反映。当论文篇名写得过大或过小时,就不能仅仅局限在篇名的范围,而应进一步在论文的摘要中选择,甚至需要在通读全文后再选取关键词。对医学期刊来说,关键词可以是研究目的、研究对象、

诊疗方法、部位或器官、区域等。举例如下:

例1.“中老年高血压患者高脉压与冠状动脉病变的相关性”一文的关键词:血压;高血压;冠心病;中年人;老年人

例2.“波姆光治疗仪治疗宫颈糜烂100例”一文的关键词:

原文为:治疗仪;宫颈糜烂;感染(宫颈糜烂;波姆光;光化学疗法)

例3.“湖北松滋地区乙型肝炎病毒感染情况调查”一文的关键词:

原文为:乙型肝炎病毒;血清标志物;病毒载量(肝炎;乙型;肝炎病毒;乙型;肝炎抗体;乙型;肝炎抗原;乙型;HBV-DNA;湖北)

2 关键词的选取原则^[1]:

2.1 宜选取有实指意义的词语

由于一些作者对论文主题内容理解不够,不太熟悉关键词的选择,或不懂关键词的组配原则,往往随意选用一些词充当关键词^[1]。诸如:“作用机制、

收稿日期 2010-10-14

作者简介 陈秋莲,女,馆员。

诊断、手术、治疗、临床应用、首发症状、疗效观察、调查分析、初步探讨、病例报告、研究、发展、方法对比、监测、检查、临床意义”。这些“万能词”部分是副主题词的范畴,不能独立存在,部分词概念抽象,没有什么信息量,并不能反映文章内容的主题,标引该词,对该文献的检索毫无帮助。举例如下:

例1.“冠心病发作胸痛患者血清酶标志物检测意义”一文的关键词:

原文为:冠心病,胸痛,血清酶标志物(肌钙蛋白I,肌红蛋白,肌酸激酶同工酶)

例2.“心肌细胞凋亡及其相关基因”一文的关键词:

原文为:心肌细胞,凋亡,基因(ICE基因,fas基因,bcl-2基因,p53,c-myc基因)

例3.“结核分支杆菌耐药基因的检测方法研究进展”一文的关键词:

原文为:结核分支杆菌,耐药,基因突变,检测(聚合酶链反应,基因芯片,DNA直接测序,DNA指纹)。

2.2 宜短不宜长

关键词之间可相互组配,但词与词之间意思大多数是孤立的,语法上毫无关联,因此用过长的句子成分来代替关键词是不适宜的,也不符合尽量用规范的检索语言来标引关键词的要求。

如:

(1)小肠结肠炎耶尔森氏菌:应为:肠炎,耶尔森氏菌。

(2)小儿神经母细胞瘤:应为:神经母细胞瘤;儿童。

(3)尺桡骨干双骨折:应为:尺骨,桡骨,骨折。

(4)晚期血吸虫病:应为:血吸虫病

而一些规范的缩略语,如“播散性血管内凝血”(DIC)、“急性呼吸窘迫综合征”(ARDS)、“酶联免疫吸附试验”(ELISA)已被公认并广泛应用,则另当别论。

3 关键词首标词的选择

首标词是表达文章主题最核心的关键词,它是数个关键词第一个标引的词。它的选择是否准确,直接关系到标引的质量并影响到医学文献的导读^[4]。首标词选择的准确、合理,揭示了文献主题最本质的内容,就能大大缩小读者查找文献的范围。

医学期刊在标引数个关键词后,往往忽视首标词的选择,而信手拈来将其随机排列的情况仍为多

见。由于篇名大都以研究对象和实验(手术)方法排列在前,故以研究对象、实验方法和诊疗手段作首标词的相当普遍。

例1:“1210例10岁以下儿童血铅检测结果分析”关键词:儿童,血铅,分析。首标词应为:血铅。

例2:“超高倍显微仪检查粪便图片诊断肠道真菌感染的临床意义”关键词:超高倍显微仪,慢性腹泻,真菌感染。首标词应为:腹泻。

例3:“抗精神病药物对精神分裂症患者血清总胆汁酸的影响”关键词:抗精神病药物,精神分裂症,总胆汁酸。首标词应为:精神分裂症。

4 关键词选取存在的问题

4.1 漏标

关键词的选择范围可来自题名、摘要、全文,可以是研究目的、研究对象、诊疗方法、部位、器官、区域。如果论文选择的关键词过少,不能完全反映文献的主题,影响该文献的检出,或对论文的主要研究目的不理解,造成论文的主要关键词未标引,这就是漏标^[2]。如:

例1:“内镜微波治疗吻合口狭窄42例”一文中,关键词:内窥镜术,微波,吻合口狭窄。漏标:食管贲门癌。

例2:“消化道溃疡幽门螺旋杆菌感染与胃酸及胃泌素的关系”一文中关键词:幽门螺旋杆菌,胃酸,胃泌素。漏标:胃溃疡,十二指肠溃疡。

例3:“前列腺液的细菌培养及药敏分析”一文中关键词:前列腺液,细菌培养,药敏试验。漏标:前列腺炎。

例4:“562份男性不育症精液常规检验结果的分析”一文的关键词:精子密度,精液液化时间,精子活动率。漏标:不育症。

4.2 错标

在医学文献中存在着大量词义有关联的医学术语,如果不理解这些术语的概念,则会混淆,造成误标。如:破伤风杆菌—破伤风,血吸虫—血吸虫病;乙型肝炎病毒—乙型肝炎;人类免疫缺陷病毒—HIV感染,阿托品—阿托品化,洋地黄—洋地黄化。

4.3 复标

医学文献中同义词、近义词应用范围较大,词意有抽象词、实意词,概念上有些词是大范围的泛指,有些词义具体属于专指。将具有上、下位关系的词同时标引,或虽字面上看词义不同,但实际表达

的主题内容是一致的词同时标引造成复标。由于选择相近词义的关键词造成概念重复,使医学论文标引的几个有限关键词多了一二个无效信息点。

例 1. “定西市 18 岁以下人群 5 种微量元素检查结果分析”一文中关键词:微量元素;铁;锌;铜;钙;镁;定西。其中铁、锌、铜、钙、镁均属微量元素,关键词“微量元素”属重复标引。

例 2. “泌尿生殖道支原体感染和耐药性分析”一文中关键词:支原体;解脲脲原体;人型支原体;药物敏感试验。其中解脲脲原体、人型支原体均属于支原体,故关键词“支原体”属于复标,应删除。

当然,关键词也并不是越多越好,选择太多造成滥标,势必重点不突出,专指性差,降低了检索效果,影响了查准率。

另外,当论文报道的内容较先进,对新出现的名词,相应的主题词暂未出现,仅有上位词时,个人建议直接使用新名词,这更有助于该文献的检出,

及查新检索相关课题时减少漏检。另外,目前研究中对部分基因,研究人员更倾向于使用英文,该基因的关键词建议亦使用英文。如 survivin 基因,部分文献翻译为存活素,但用“存活素”检索时,检出的相关文献量远远少于用“survivin”,影响该文献的推广及查新检索,如 BCL-2 基因,没有中文名称,只有上位主题词,该词的关键词标引应写“BCL-2 基因”。

综上所述,论文关键词的选取在论文写作及论文编辑中是不可忽视的重要环节,遵照一定的原则选取准确的关键词,有助于文献的检出及推广。

参考文献:

- [1] 徐鸿飞, 缪宏建. 医学期刊关键词选择的多维构想[J]. 南通医学院学报, 1996, 16(4): 634-636.
- [2] 胡美君. 浅谈查新工作中如何提高查全率和查准率[J]. 浙江高校图书馆情报工作, 2005, (4): 40-41.

(编辑 朱爱瑜)

(上接第 53 页)人电脑(L5),有的图书馆只填报几台,不免让人存疑,经了解后得知,是错将个人电脑理解为“归个人使用的电脑”了,如果将 L5 原有注释“年末在用的所有个人电脑数”改为“年末在用的所有 PC 机以及笔记本电脑”,或许可以使该统计项更容易理解。

总之,任何一个统计指标体系,都离不开对统计指标进行说明的文字和文件,在制定说明性文件时,既要保持语言的准确性和学术性,还要力争使之易读、易懂。

4 结论

教育部事实数据库是我国高校图书馆统计工作的规范和标准,要使其在今后的工作中发挥更大的作用,必须做到以下三点:一是指标系统的科学化,填报说明的详尽化;二要根据图书馆事业的发展,及时更新指标体系以及相关说明;此外,对统计人员的培训,也是做好统计工作的重要环节。

参考文献:

- [1] 姜晓等. 高等学校图书馆统计[M]. 成都: 成都科技大学出版社, 1995: 17.
- [2] 教育部高校图书馆事实数据库[DB/OL]. <http://162.105.140.111/calis/default.asp>.
- [3] 中华人民共和国高等教育法[EB/OL]. http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/moe_619/200407/1311.html.
- [4] 为何在职研究生考试无学历证书只有学位证?[EB/OL]. <http://www.onjobedu.com/html/hdwd/kaoshi/50768.html>.
- [5] Google Analytics [EB/OL]. <http://www.google.cn/analytics/>.
- [6] 教育部高等学校图书馆情报工作指导委员会, 高等教育文献保障系统(CALIS)管理中心. 高等学校图书馆数字资源计量指南(2007 年)[EB/OL]. http://162.05.140.111/tugon-gwei/info/index.asp?strCode=specification_63.
- [7] 高教资源. 中国高等学校教师网[EB/OL]. <http://www.ccf.edu.cn/zhiyuan/>.
- [8] 顾健. 美国教育统计中心的大学图书馆统计[J]. 大学图书馆学报, 2008, (2): 15-21.

(编辑 朱爱瑜)