

# 信息传播视角下学术论文关键词标引方法与策略

■张红霞 冀伦文 贾丽红 王晶冰

收稿日期:2018-08-13

修回日期:2018-11-14

太原理工大学期刊中心,山西省太原市迎泽西大街79号 030024

**摘要** 【目的】探寻实用有效的关键词标引方法,以期解决关键词标引不正确的问题。【方法】参考 GB/T 3860—2009《文献主题标引规则》,结合中国知网的中国学术期刊全文数据库,从信息传播视角探讨关键词标引方法的科学性与合理性。【结果】提出关键词标引的步骤:通过主题分析,用研究对象、研究方法、结果结论、其他重要信息4个要素来确定主题概念;采用组配、上位词等方式把主题概念转换为关键词;借助学术期刊全文数据库对所选关键词进行辅助性审核和修改。【结论】科学、合理地标引关键词有助于提升文献利用率和期刊学术影响力。

**关键词** 学术论文;关键词;标引;信息传播;数据库;主题概念;学术影响力

DOI: 10.11946/cjstp.201808130718

关键词是为了便于文献检索而选取的、能够反映论文主题的词语或术语。一般每篇论文标注3~8个关键词,并尽量采用《汉语主题词表》等提供的规范词,未被各类词表收录的新学科、新技术中的重要术语和地区、人物、文献、产品及重要数据名称,也可作为关键词<sup>[1]</sup>。

随着科技的飞速发展,文献量急剧增长,依靠各种学术文献数据库检索文献已成为人们获取学术信息的主要方式。如果说出版是信息传播的第一步,那么数据库检索中的“被发现”就是信息传播的第二步,也是传播链上最关键的一环;若此环松动、掉链,文献“被捞出”的概率将很小,甚至石沉大海。论文“被发现”的途径主要包括主题检索、关键词检索、摘要检索、标题检索、作者检索等,其中关键词检索是用户最常用的文献检索方式之一,只有用户检索的关键词与某篇论文的关键词一致时,这篇论文才会被检出。因此,关键词作为论文信息传播的源头,在论文的检索过程中担负着文献“被正确、恰当、充分地检出”的重任。学术期刊编辑作为学术信息的传播者<sup>[2-3]</sup>,承担着信息选择、加工、传播的使命,要特别重视论文关键词的标引工作。关键词标引质量的高低直接关系着学术文献的有效传播和充分利用,从而影响着期刊的学术影响力。

自20世纪80年代起,学者们就持续不断地进行着学术论文关键词标引的研究,发现学术期刊论

文普遍存在着关键词标引不当的问题,主要表现为:所选关键词不能准确、全面地反映论文主要内容,存在主题信息遗漏、不清晰、不准确的问题<sup>[4]</sup>;采用无检索价值的泛义词作为关键词;把标题随意拆分成几个词作为关键词<sup>[5]</sup>;为了表达某一确切含义,把关键词写成多个内容全面的短语<sup>[6]</sup>;关键词(包括英文关键词)表述不规范<sup>[7-8]</sup>;缩略词使用不当<sup>[9]</sup>等。关于关键词的标引,学者们也给出了一些方法和建议,例如:应从研究目的、研究对象、研究方法、研究结果等方面入手,所选出的关键词才能准确体现论文的中心内容<sup>[10-11]</sup>;关键词的逻辑组合应能鲜明直观地体现论文主题内容<sup>[12]</sup>;关键词的标引步骤包括先对论文进行主题分析,然后确定若干主题概念,再把主题概念转换为关键词<sup>[4,13-14]</sup>;依照有关国家标准,关键词标引要遵循专指性原则、组配原则和上位词标引原则<sup>[10,14]</sup>;要多琢磨用户检索文献时的思路和可能采用的词汇,提高论文的被检率和利用率<sup>[5,12]</sup>。这些研究成果为关键词的正确标引提供了很好的指导与帮助,但鲜有学者对所提出的标引方法和选词原则进行深入探讨,仅以几个简单例子对所给出的方法予以说明,可操作性不强。因此,有必要进一步探讨关键词的标引方法。

通过对国家标准 GB/T 3860—2009《文献主题标引规则》的研读、分析<sup>[15]</sup>,笔者发现该国家标准遵循对文献进行主题标引的一般规律,主要是针对有主题

**作者简介:**张红霞(ORCID:0000-0002-4524-6683),学士,副编审,E-mail:1749891605@qq.com;冀伦文,硕士,副教授;贾丽红,学士,编辑;王晶冰,硕士,副教授。

词表可查的受控标引,可为关键词标引提供参考。因此,本研究参考国标 GB/T 3860—2009,结合中国知网的中国学术期刊全文数据库,从信息传播角度探讨如何科学、合理地标引关键词,提出比较系统的关键词标引方法,并以一篇出版方面的论文关键词修改实例来说明所提方法的有效性与实用性。

## 1 关键词标引方法与步骤

本研究以论文被正确、恰当、充分检出为目标导

向、围绕信息传播效果提出关键词标引方法。此法与文献[4,13-14]的提法有些类似,也分为三大步骤(表1):(1)通过分析论文主题确定若干个主题概念;(2)将主题概念转换成关键词;(3)借助文献数据库对所选关键词进行辅助性审核与修改。标引过程合理有序,每一步不仅有明确的目标,还有相应的操作方法,如确定主题概念的“四要素法”、主题概念转换为关键词的三种方式、利用数据库辅助性审核标引结果的方法等。

表1 关键词标引步骤说明

步骤	任务	目标
第1步	对论文进行主题分析,用“四要素”法确定出若干主题概念	提炼出的主题概念应能够准确、全面地反映论文的核心信息
第2步	采用组配、上位词等方式把主题概念转换成关键词,或直接将主题概念词作为关键词	关键词的组合应与主题概念的内涵基本一致;关键词应为规范的术语或词语,表达形式较固定,有一定的使用频率,以便于对论文的检索
第3步	借助学术期刊全文数据库对所选的关键词进行辅助性审核与修改	解决所选关键词可能存在的突出问题(主要是术语不规范问题及某些词在数据库中使用时频率极低的问题),进一步保障论文的查全率与检准率

### 1.1 通过分析论文主题确定主题概念

学术论文主题是论文所论述与研究的问题,即论文的核心信息和最主要的内容。论文题名往往只能在一定程度上反映论文主题,如研究对象、研究方法等,全面、准确的主题信息还需进一步研究分析;而这对于具有相关专业背景、可获得作者帮助以及对论文已进行数遍加工与校对的编辑而言,找出论文主题并不难。主题概念是指描述论文主题基本要素的概念。笔者认为,既然论文摘要集中反映论文的主要内容或者主题信息,在确定论文主题概念时可借鉴摘要的四要素法(目的、方法、结果、结论);同时又考虑到主题概念的词性(名词性词语或相当于名词的术语)以及用户检索的习惯与需求,最终

确定表达主题概念的四要素为研究对象、研究方法、结果结论和其他重要信息。这与文献[11]提出的主题四要素(研究目的、研究类别、研究方法、研究结果)不同。表2就本研究提出的主题四要素的内涵分别予以解释,并给予实例说明。

在确定主题概念时要遵循三个原则:

- (1) 客观性。客观分析论文实际具有的主题,这就要求编辑能把握住论文的核心信息。
- (2) 一致性。主题概念所反映主题的全面性与专指性应与论文实际具有的主题全面性与专指性基本一致。
- (3) 实用性。充分考虑检索系统目标与用户的需求,选定论文中有检索意义的主题概念。

表2 主题概念的四要素内涵及实例说明

主题四要素	各要素内涵	实例说明
研究对象 (研究什么?)	论文的研究对象,可以是问题、产品、材料、设备、事件、现象等,以及研究对象的某一方面(如结构、成分、材料、性质等)	题名为《中国科技期刊刊名字体设计研究》的论文,研究对象为“科技期刊刊名字体设计”
研究方法 (如何研究?)	进行研究所用的原理、理论、条件、材料、工艺、结构、手段、装备、程序等	题名为《中国科技期刊刊名字体设计研究》的论文,研究方法为“设计原则”“期刊定位”“期刊品牌”
结果结论 (研究得出什么?)	论文的创新点,或得到的新理论、新观点、新产品、新技术、新材料、重要数据等。有时一篇论文的创新就体现在研究对象、研究方法的创新上,所以“结果结论”要素有时可以没有	题名为《Publons管理模式对我国同行评审制度的启示》的论文,结果结论为“开放同行评审”
其他重要信息 (还有什么有检索价值的重要信息?)	具有检索价值的空间(地区)、时间(时代)、人物、文献、事件、机构等的名称。另外要考虑文献是否具有潜在用途	题名为《昭通褐煤半焦气化特性的研究》的论文,应把地名“昭通”列为其他重要信息

### 1.2 将主题概念转换成关键词

#### 1.2.1 主题概念转换成关键词的必要性及转换方式

主题概念转换为关键词有3种方式,如表3所

示。其中主题概念词即代表主题概念的词语,可以是一个规范的、形式较固定的词语或术语,也可以是复合程度较高、形式不固定的词组。

许多作者和编辑不理解关键词组配表达主题概念等方式,往往将关键词写成“内容全面”的短语;此时关键词的限定词增多、复合程度升高、专指性增强,所检出的文献数量减少,从而严重影响了文献的被检率和利用率。对数据库用户而言,其检索目标是就某一主题获取尽可能全面的相关文献,关心的是查全率,而不是某一篇或几篇文

献。因此,在主题概念形式不固定、复合程度较高的情况下,有必要采用组配方式将主题概念转换成2个及以上的关键词。用户可通过使用关键词的组合或者采用高级检索方式进行检索,以缩小检索范围,提高检准率。总之,不管采用哪种方式,均是为了便于论文能得到充分、合理、有效的检索。

表3 主题概念转换关键词的3种方式

转换方式	适用情形	举例说明
主题概念直接作为关键词	主题概念为规范的、形式较固定的词语或术语	科技期刊、人类命运共同体、以人为本、表面活性剂、动力屈曲等可直接作为关键词
主题概念转换为2个或2个以上的关键词(即组配的方式)	主题概念为形式不固定、复合程度较高的词语	相移布拉格光栅→相移光栅+布拉格光栅(交叉组配);科技期刊刊名字体设计→科技期刊+刊名字体设计(方面组配)
选用主题概念词的上位词作为关键词,必要时该主题概念词也可作为关键词	主题概念属于新术语,代表研究领域内的新观点、新理念、新成果,但该术语尚未得到普及	学术运营→运营+学术运营。“学术运营”专指将学术与市场进行融合的运用,与普通的运营有明显不同。这样既能保障论文的查全率,又能体现论文的创新点

注:“→”表示转换,“→”前为主题概念词,“→”后为关键词。

### 1.2.2 主题概念转换成关键词的注意事项

(1) 可以通过查词表来确定专业术语表达的规范性。此类词表可以是《汉语主题词表》、某学科的专业术语词典、全国科学技术名词审定委员会颁布的科学技术名词或医学 MeSH 词典<sup>[16]</sup>。

(2) 实际上许多学科没有专业术语词典,并且专业词典的更新速度远远赶不上新术语的涌现速度。这种情形下有2种处理方式:①若该词是同行认可的、形式较固定的、规范的或使用频率高的术语,则直接作为关键词;②若该词的复合程度较高、表达形式不固定且可拆分,则采用关键词组配的方式。

(3) 主题概念转换成关键词时,要注意选用有检索意义的词,这类关键词可分为具有独立检索意义的词和具有辅助检索意义的词2种,如国际证券市场、盾构隧道、金属有机骨架材料,及数值模拟、综述、算法、调研等。而研究、探讨、思考、分析等没有检索意义的词应避免用作关键词。

### 1.3 借助数据库审核和修改关键词

关键词的规范性、使用频率对查全率的影响及其组配方式、表达方式的合理性都需要进一步审核分析,而数据库便是最佳的辅助媒介。

数据库辅助审核、修改关键词步骤如下:(1)选择关键词作为检索条件,将选好的关键词输入检索框;(2)通过检出文献数判断关键词的使用频率、规范性及适用性;(3)依据判断结果对所选关键词进行修改。虽然根据检出结果不一定能准确判断该词的适用性,但如表1所述,可以解决所选关键词可能存在

的突出问题,进一步保障论文的查全率与检准率。

为了说明所提方法的有效性,本研究以某些论文中的关键词作为检索词,利用关键词检索方式进行检索,得出检出结果及关键词修改意见(表4)。

(1) 可利用中国学术期刊全文数据库检验关键词表达的规范性。编辑加工学术论文时,经常会遇到专业术语前后表述不一致的情形,此时可以通过专业词典、数据库等判断专业术语的规范性、通用性及适用性。例如表4中的“MOF”与“MOFs”、“蒸气渗透”与“蒸汽渗透”、“金属有机框架材料”与“金属有机骨架材料”,通过数据库审核发现每组均是不规范表达与规范表达的问题,只有后者才符合原论文的本意。

(2) 编辑在对专业术语不熟悉的情况下,可以通过检出文献数得出该词的使用频率,进而决定是直接主题概念词作为关键词,还是采用组配的方式来表达。例如表4中“燃气轮机冷却系统”和“深埋硬岩隧道”使用频率极低,因此不适合直接用作关键词,必须采用关键词组配的方式。

(3) 如果某一关键词的检出文献数特别少,应视具体情况进行处理,可采用将上位词作为关键词、原词与上位词同时作为关键词这2种处理方式。如用“单齿配体”代替“非手性单齿配体”、用“封面设计”“字体设计”代替“刊名字体设计”。需要注意的是,部分关键词所代表的研究方向为学科前沿,虽然检出文献数较少,但仍应用作关键词,如“锂硒电池”一词,检出的文献只有7条,但仍应用作关键

词,并增加其上位词“锂离子电池”作为关键词。

(4) 代表研究对象、研究方法的关键词可以称之为核心关键词,对这些词的审核要格外慎重。如

果这些词作为检索词且检出文献数特别少时,要考虑换词,如采用其上位词或采用两个使用频率高的关键词组配的方式来提升论文的被检索率。

表4 以某些关键词作为检索词在数据库中检索的结果及关键词修改意见

检索词	检出文献数	修改意见
非手性单齿配体	1	研究方向偏冷,可选用其上位词:单齿配体(检出文献数:25)
锂硒电池	7	锂硒电池是一种新型的锂离子电池,且检出文献的发表时间均为近三年。因此保留该关键词,并加上其上位词“锂离子电池”作为关键词
刊名字体设计	0	其相关关键词的检出文献数:刊名字体(1);字体设计(1233);封面设计(1706)。因此采用上位词“封面设计”和“字体设计”的组配代替“刊名字体设计”作为关键词
燃气轮机冷却系统	0	采用组配的方式改为“燃气轮机”和“冷却系统”两个关键词。检出文献数:燃气轮机(7652);冷却系统(4227)
深埋硬岩隧道	1	采用组配的方式改为“深埋隧道”和“硬岩隧道”两个关键词。检出文献数:深埋隧道(290);硬岩隧道(47)
MOFs	502	MOFs为“金属有机骨架材料”英文名称简称,符合原文献的意思,不做改动。MOFs可用于关键词的英文表达。
MOF	1284	MOF为元对象机制、多器官功能衰竭等词语的简称,不符合原文意思,不采纳
蒸汽渗透	136	符合原文意思,不改动
蒸气渗透	28	该词为“蒸汽渗透”的不规范写法,不采纳
金属有机骨架材料	231	符合原文意思,不改动
金属有机框架材料	61	“金属有机框架材料”与“金属有机骨架材料”为同义词,在同一篇论文中出现这两种写法。通过数据库审核,发现“有机骨架材料”的使用更普遍、更规范。因此不使用“金属有机框架材料”这种写法。

注:所用的数据库为中国知网的“中国学术期刊全文数据库”,检索方式为关键词字段检索;检索日期为2018-11-09。

特别需要说明的是,表4中的例子不是个例,在审核关键词时会经常遇到这些情形,此时只有经过关键词审核和修改,才能对所选关键词的合理性有一定把握,故十分有必要利用数据库对关键词进行辅助性审核与修改。

## 2 实例说明

在实际工作中,编辑所面对的是作者已标出的论文关键词。因此在运用本方法修改关键词时,需要在把握论文的核心内容并与作者充分沟通的前提

下,以已标出的关键词为参考来重新标引关键词。

为了便于编辑同仁对该方法的理解,笔者专门选取一篇出版类论文作为实例,运用本研究的关键词标引方法对论文原有关键词进行修改(表5)。该文题为《上海市英文科技期刊出版现状分析与思考》,主要内容是从期刊的基本情况、期刊采编系统、期刊网站建设、期刊国际合作出版、期刊出版模式等方面对上海市英文科技期刊的出版现状进行调研,得出亟需提升英文科技期刊数量和质量、期刊专业化程度、期刊网站建设质量与期刊国际影响力的结论。

表5 运用本研究方法修改论文关键词的实例

标引每步的结果	研究对象	研究方法	结果结论	其他重要信息
由第1步得到: 主题概念	英文科技期刊出版现状	调研; 采编系统; 期刊网站; 国际合作出版; 期刊出版模式	国际影响力	上海市
由第2步得到: 主题概念转换后的关键词	英文科技期刊; 出版现状	调研; 采编系统; 期刊网站; 国际合作; 出版模式	国际影响力	上海市
由第3步得到: 审稿后的关键词	英文科技期刊(148); 出版现状(220)	采编系统(675); 期刊网站(476); 国际合作(出版类33)	国际影响力(出版类243)	上海(139590)
原文关键词对照	英文科技期刊(148); 出版现状分析(1)		国际影响力(出版类243)	

注:括号内的数字为在关键词检索方式下,以该关键词作为检索词的检索结果(检出文献数);检索日期为2018-11-09。

通过对实例关键词的标引,可以得出以下结论。

(1) 该方法具有可操作性。作者和编辑有章可依、心中有谱,就不会“任性”地选择关键词,不易产生信息遗漏、标引不准确的问题。

(2) 修后关键词个数较原文关键词多了4个,更全面地反映论文主题内容,便于数据库用户从多个角度检索到该论文,从而充分发挥论文的学术价值。Uddin等<sup>[17]</sup>研究论文被引频次与关键词个数之间的关系,发现在不考虑论文学术质量的前提下,论文被引频次与关键词个数呈正相关,即论文关键词个数越多,被引频次越高。因此,为使文献得到充分的利用,关键词个数偏多为好( $\leq 8$ )。

(3) 由主题概念转换成关键词时,要充分考虑词语的规范性和使用频率。如国际合作出版(6)→国际合作(出版类33)、期刊出版模式(1)→出版模式(538)、上海市(79642)→上海(139590)、出版现状分析(1)→出版现状(220)(括号内数字为以该词作为检索词的检出文献数,检索日期为2018-11-09)。建议有关部门尽快建设我国的主题词电子词库<sup>[18-19]</sup>,便于编辑及用户检索查询,提高关键词的标引精度和数据库的检索效率。

(4) 对转换后的关键词进行优化与精选。最终确定的关键词比转换后的关键词少了“调研”和“出版模式”。删除原因在于“调研”是该文的研究方法,意义较泛;“出版模式”在原文中探讨较少。

(5) 有针对性地审核关键词,不是每一个词都需要审核。不需要审核熟知的规范词和专业术语,重点审核那些无法判断其规范性和使用频率的关键词。

### 3 结语

本研究提出一个比较系统的关键词标引方法:从研究对象、研究方法、结果结论、其他重要信息4个方面去确定论文的主题概念,以准确、全面地反映论文的主题信息;采用组配、上位词等方式把主题概念转换为关键词或者直接把主题概念词作为关键词;并借助文献数据库对所选关键词进行辅助性审核和修改。

通过上述研究,笔者对编辑关键词标引工作及我国文献数据库的发展提出以下2点建议:(1)关键词标引质量的高低直接影响学术文献的有效传播和充分利用,因此要高度重视关键词的

标引工作,建议把关键词标引质量控制作为科技期刊编辑规范化的一项内容,并加强关键词标引方面的培训。(2)为大幅提升文献的传播效果和利用价值,我国学术期刊文献数据库需要加快主题词库建设步伐,向优秀数据库学习。如国外的Web of Science数据库中每篇文献的关键词除文献原有的关键词外,还添加了计算机自动抽取的关键词;EI数据库有自己的主题词(受控词)库,每篇文献的关键词除文献原有的关键词外,还标出一定数量的受控关键词,进一步提升了文献被发现、被利用的概率。

作为关键词标引方面的探讨,本研究希望能给予从事、研究关键词标引工作的人员提供一些有益的帮助和思路,同时也欢迎各位对研究的不足之处批评指正。

### 参考文献

- [1] 新闻出版总署科技发展司,新闻出版总署图书出版管理司.作者编辑常用标准及规范[M].3版.北京:中国标准出版社,2012:514-524,557-568.
- [2] 刘杨,赵大良,葛赵青.学术期刊信息的传播模式[J].编辑学报,2005,17(6):410-412.
- [3] 张小强,张莘,吕赛英,等.从信息传播角度看科技期刊编辑出版过程及其优化[J].编辑学报,2007,19(3):161-163.
- [4] 倪向阳,马永军.科技期刊的关键词标引质量亟待提高[J].编辑学报,2011,23(4):305-307.
- [5] 程智强,江洪涛.科技论文关键词标引应琢磨的几个问题[J].编辑学报,2014,26(1):40-42.
- [6] 王立名.科学技术期刊编辑教程[M].北京:人民军医出版社,1997:50.
- [7] 汪新红.科技期刊论文网络传播的编辑特点[J].中国出版,2011(2):58-60.
- [8] 于双平,王文武,姜晓舜.中英文关键词不一致现象统计分析[J].编辑学报,2009,21(6):498-499.
- [9] 闫杏丽,张淑艳,张镭.关键词中缩略词的标引现状及建议[J].编辑学报,2013,25(4):349-352.
- [10] 曹素文.护理论文关键词的标引[J].护理实践与研究,2006,3(1):59-61.
- [11] 张建蓉,陈燕.学术论文中关键词标引的常见问题剖析[J].编辑学报,2003,15(2):104-105.
- [12] 李力.关键词在学术论文中的合理使用[J].情报探索,2003(3):3-4.
- [13] 王昌度,熊云,徐金龙,等.科技期刊论文关键词标引的问题与对策[J].编辑学报,2003,15(5):349-351.
- [14] 张积玉.论社会科学论文关键词及其标引规范[J].吉林大学社会科学学报,1993(6):93-94.
- [15] 中华人民共和国国家质量监督检验检疫总局,中国国家标准化管理委员会.文献主题标引规则:GB/T 3860—2009[S].北京:中国标准出版社,2010.
- [16] 杨晨晨.利用万方医学网辅助医学编辑查找标注主题词作为

关键词的方法[J]. 江苏科技信息, 2013(8):22-24.

- [17] Uddin S, Khan A. The impact of author-selected keywords on citation counts[J]. *Journal of Informetrics*, 2016, 10: 1166-1177.
- [18] 杜香莉, 王立宏, 罗红彬. 我国期刊全文数据库关键词规范化问题探讨[J]. 中国科技期刊研究, 2007, 18(4): 612-614.
- [19] 刘大乾. SCI关于关键词的一般选取准则及词义库建设[J]. 中国科技期刊研究, 2007, 18(6): 1073-1074.

#### 作者贡献声明:

张红霞: 提出研究思路, 撰写、修改论文;  
冀伦文: 调研与整理文献;  
贾丽红: 参与论文修订;  
王晶冰: 参与论文最终版本修订。

## Keyword indexing method and strategy of academic paper from the perspective of information dissemination

ZHANG Hongxia, JI Lunwen, JIA Lihong, WANG Jingbing

Journal Center, Taiyuan University of Technology, 79 Yingzexi Street, Taiyuan 030024, China

**Abstract:** [Purposes] In order to solve the problem of incorrect keyword indexing, this paper aims at exploring the simple and effective keyword indexing method. [Methods] Referring to GB/T 3860—2009 *Guidelines for Determining the Subjects and Selecting the Descriptors* and combining the Chinese Journal Full Text Database in CNKI, we analyzed the scientificity and rationality of keyword indexing method from the perspective of information dissemination. [Findings] The procedures of keyword indexing method are as follows: determining subject concepts through subject analysis, which are determined by research object, research method, results and conclusions, and other important information; then transforming subject concepts into keywords by the means of coordination and broader term; at last examining and revising keywords by the aid of the database. [Conclusions] The scientific and rational keyword indexing method will improve the utilization rate of literature and academic influence of journals.

**Keywords:** Academic paper; Keyword; Indexing; Information dissemination; Database; Subject concept; Academic influence

(本文责编:刘晶晶)