

研究简报

水产科技文献中使用关键词的若干问题

SOME ISSUES ON USING KEYWORDS TO SHOW SUBJECTS OF FISHERY SCIENTIFIC LITERATURES

李竹青 王民生

(中国水产科学研究院, 北京 100039)

Li Zhuqing Wang Minsheng

(Chinese Academy of Fishery Sciences, Beijing 100039)

关键词 水产科技文献, 关键词, 标引方法

KEYWORDS Fishery technological document, Keyword, Indexing

近年来,许多水产科技刊物为了建立自己的年度索引,都做了关键词的标引(给文献赋予检索标识);广大科技工作者由于情报意识的提高也都自觉地对自己的论文、报告提供了关键词。这反映了水产科学技术水平以及科技人员素质的提高。但是,水产系统使用主题法(采用主题词作检索标识的方法)加工、处理文献的工作才刚刚起步,一些同志对此还比较陌生。

作者对近年 200 多期水产科技出版物中论文、报告所标的关键词做过调查,发现有些问题尚需明确。本文就这些问题从正面提出意见,供广大科技工作者参考。

关键词及其意义

关键词是主题词的一种,它是从文献(标题名、章节名以及摘要、正文)中直接抽取出来的、具有检索意义的,即表示文献关键(主题)内容、并且可作为检索入口的主题词。

目前,有些科技出版物把所使用的标引词称为“主题词”,这是不准确的。因为这些词是直接从文献中抽取出来的,确切地说应该称为“关键词”。

给每一篇论文、报告标出关键词,提出检索标识,目的是为了区分和识别文献,作为有序存贮和检索的依据。无检索标识的文献既无法进入检索系统,也无法进行检索利用。广大科技工作者为自己的论文、报告正确地标出关键词,这是将自己的成果提供给广大用户

收稿日期:1992-09-30。

利用的必不可少的工作。关键词是论文、报告的重要组成部分,故标引关键词是提高出版物利用价值的重要措施。

另外,科技工作者要利用主题法查找文献,首先要根据自己所需要的文献内容范围,提出关键词,或者再由各检索系统将关键词转换成本系统所用的叙词,从而实现检索文献的目的。所以,关键词也是科技工作者查找文献、资料,寻求新知识的一把钥匙。科技工作者掌握主题标引和检索技术,可以得到事半功倍的效果。

文献主题的确定

标关键词一般分两步。首先是确定文献的主题,掌握文献的主题要素;第二步是将主题要素用关键词表达出来。这两步有时很难区分,但缺一不可。

一般来讲,文献主题有单一主题和多主题之分。所谓单一主题,是指只有一个中心内容,例如《中国对虾的人工授精》,即属于一个主题的文献(假设该文献的标题表达了其主题,下同)。所谓多主题,即文献包含两个或两个以上的中心内容,例如《江苏省的海洋渔业和淡水渔业》,即属于有两个主题的文献。

文献的主题即使只有一个,它也会由几个主题要素所组成。例如对象、属性、条件、方法、过程、结果等,它反映着主题的各个方面。例如,《中国渔业用水的水质标准》这篇文献,其主题要素便是中国(条件)、渔业用水(对象)和水质标准(属性)。体现一个主题的最关键的主题要素,叫做主题中心。上述的例子主要是论述“水质标准”(水质的各种指标和要求),所以,这篇文献的主题中心是“水质标准”。这里的“中国”、“渔业用水”是指明该标准的使用范围,是主题的条件和对象面,是一般主题要素。

一般的文献标题与主题的关系比较复杂,不是只从文献标题中抽出几个词就可以反映主题的。据作者统计,水产文献中标题不反映主题,或者不完全反映主题的占 15—65%。例如《我区渔业之管见》和《LRA—851 型集成化定位仪》这两篇文献,它们的主题就很难从标题中看出。所以,要正确地标引关键词,必须首先细心地分析文献的内容及其重点,对标题要进行分解分析,对重点内容进行概括分析,找出表达主题的全部关键词。

当然,所抽出的关键词也不是越多越好,以反映出文献的主题及主题的主要方面为度,一般 3—5 个关键词就可以了,多了反而淹没了重点。

确定主题时应抓住以下重点:一是论文所要研究、探讨的主要对象和主要目的;二是处理主要对象所采用的工艺、方法、设备或者材料、属性等;三是其他有检索意义的重要线索,如地名、人名和年代等。

在确定主题的过程中,容易出现的毛病是:1. 主题概念提炼不全;2. 主题概念提炼过多;3. 主题概念提炼错误等。现以《中国水产文摘》910226 号文摘为例,试分析三种易出现的毛病。

910226 关于海带夏苗脱落的研究/丛沂滋(烟台市水产研究所)//齐鲁渔业。—1990,(4)。—8—9

海带夏苗培养技术的创造发明,在海带养殖生产中发挥了重大作用。但近几年来,育苗后期出现的幼苗脱落问题,

已成为最突出问题。本文研究了幼苗脱落的规律,分析了脱苗原因,提出防治脱苗的技术措施:(1)培育 1—2cm 中等大小的幼苗。(2)扩大供水设施。(3)认真处理附着基。(4)附着密度应控制在 20—30 个/10×10。(5)幼苗要求在风后出库(薛鸿瀛)

这篇文献的主题是:海带夏苗脱落的规律及其防治脱落的措施。关键词应标引为:海带,夏苗脱落,防治脱苗。

1. 如果关键词标为“海带,夏苗脱落”时,就没有完全反映出文献主题,漏掉了“防治夏苗脱落”的主题面,犯了“主题概念提炼不全”的毛病。

2. 如果关键词标为“海带,夏苗培养,幼苗脱落,脱落原因,防治脱苗,供水设备,附着密度”等时,属于冗余标引,犯了“主题概念提炼过多”的毛病。

3. 如果关键词标为“海带,夏苗培养,附着基”等时,便曲解了文献主题,把主题曲解成为“海带夏苗的培养方法”,犯了“主题概念提炼错误”的毛病。

另外,在确定主题的过程中,还容易出现的一个毛病就是,误把表现文献主题的形式当成了文献的主题。例如,《国外渔业概况》(书)的主题是:国外主要渔业国家的渔业生产和渔业科技水平。关键词应标为“国外渔业,渔业生产,渔业科技”;如果标为“国外渔业,概况”就不妥当了。所谓“概况”,只不过是一种描述主题的形式,而不是主题本身。

关键词的确定

标出的关键词不仅要反映文献的主题,而且可以作为检索的入口词(可以通过该词查找到文献),使每个词都具有实际的检索意义。例如《利用计算机编制水产科学叙词表》这篇文献,其中“计算机”、“水产科学”、“叙词表”三个词能描述文献的主题,是关键性的词,并可作为检索的入口。而“利用”、“编制”两个词,虽然也能帮助描述文献主题,但不起关键作用,同时,它也不能作为检索的入口词,故不能选作关键词。理由很简单,读者要查找利用计算机编制水产科学叙词表方面的文献,一般是通过“计算机”、“水产科学”、“叙词表”等词查找,不可能通过“利用”、“编制”等词来查找。

同样,象“研究”、“试验”、“经验”、“应用”、“推广”、“现状”、“进展”、“理论”、“综述”等普通词语,都不能作为查找水产文献的入口,不能作为水产文献的关键词。

从词的类型讲,能作关键词的主要是一些名词和名词性词组,一般地说,它代表着水产业和水产科学的研究和工作对象、使用的物资,如专业、工种、学科、专题的名称,生物、工具、材料的名称以及有关的地理、年代等专名等等。不能选作关键词的词是那些作检索无实际意义的词,例如,1. 冠词;2. 介词;3. 连词;4. 感叹词;5. 代词;6. 某些副词;7. 某些形容词;8. 某些名词(如理论、报告、试验、学习、考察等);9. 某些动词(连系动词、情态动词、助动词)等等。

关键词是根据它对主题的表达、检索的需要以及科学概念的表达而确定的。因此,对词组也要根据一定原则来划分。划分的原则如下。

1. 非固定性词组或临时组成的上下文,应划分为单元词或固定词组。例如:“湖泊养

鱼拦鱼技术”这个词组,可划分为“湖泊养鱼”、“拦鱼技术”的形式选取;“黄海渔业资源”这个词组,可划分为“黄海”、“渔业资源”的形式选取。

2. 凡是本学科领域中通用的、有固定意义的词组,应以词组形式选取。例如:“资源分布”,不能划分为“资源”、“分布”选用;“好氧微生物”,不能分为“好氧”、“微生物”选用。

参 考 文 献

- [1] 张琪玉,1983. 情报检索语言,214—224. 武汉大学出版社。
- [2] 石立坚,1988. 专名与术语. 自然科学术语研究,2: 26—30。
- [3] 钱起霖,1985. 汉语主题词表标引手册,18—23,69—82. 科学技术文献出版社。
- [4] 傅兰生,1987. 主题标引与索引技术,17—23. 科学技术文献出版社。