

# 提高文献检索中查全率与查准率的方法探讨

欧丽清

(广西团校图书馆, 广西 南宁 530022)

**摘 要:** 针对衡量文献信息检索效果的主要技术指标查全率与查准率, 提出了以灵活运用算符为基础的提高查全率和查准率的文献信息检索方法。文献检索中查全率和查准率对检索人员来说, 具有非常重要的影响。查全率如果达不到标准, 检索者就查不到想要的信息, 从而影响工作的正常进度; 查准率如果达不到要求, 那么检索者得到的信息就会具有很强的模糊性, 而得到的信息也没有太大的用处。在以后的工作中, 查全率和查准率必须得到全面提升, 并且在客观上对检索者的工作产生积极影响。

**关键词:** 算符; 查全率; 查准率; 文献

中图分类号: G250

文献标识码: A

文章编码: 1672-6251(2014)04-0105-03

## Study on the Methods of Improving Recall Ratio and Precision Ratio in Document Retrieval

OU Liqing

(Guangxi Youth League School, Guangxi Nanning 530022)

**Abstract:** Recall ratio and precision ratio, the key technical indexes that measure the effectiveness of document retrieval, have a significant impact on searchers. The searchers cannot find the wanted information with substandard recall ratio, which will affect the normal work schedule, while the information obtained in document retrieval may have a strong ambiguity and be useless with substandard precision ratio. Recall ratio and precision ratio need to be fully upgraded with a positive impact on the work of searchers. This paper proposed flexible operator-based methods for improving recall and precision ratio in document retrieval.

**Key words:** operator; recall ratio; precision ratio; document

就现阶段的工作来说, 要想实现查全率和查准率的双重提升, 需要在多个方面提升。比如说布尔逻辑的能力、截词符、位置逻辑运算符的正确使用, 提升文献信息检索中的查全率与查准率。文献检索能够帮助检索者获得较多的信息, 同时根据检索者的需求及反馈回来的信息, 进一步优化查全率和查准率。在以后的工作当中, 首先要对目前的查全率和查准率进行了解, 之后要对检索者遇到的问题进行了解, 最后综合两者的信息和交集, 通过针对性的方式提升文献检索的查全率和查准率。

### 1 算符介绍

#### 1.1 布尔逻辑算符

布尔逻辑运算符用来表示两个检索词之间的逻辑关系, 常用的有三种: 逻辑“与”、逻辑“或”、逻辑“非”。布尔逻辑运算符在文件检索当中, 具有非常重

要的作用。第一种运算符是用于概念较差或者限定关系的组配, 通过逻辑“与”连接的两个检索词必须同时出现在检索结果中才能满足相应的条件。在这种情况下, 检索者需要通过清晰的词汇进行检索, 否则很难得到想要的资料和数据, 但是这种运算符可以精确检索, 减少了模糊检索带来的问题; 第二种运算符是具有概念并列关系的一种组配, 逻辑“或”连接的两个检索词中任意一个出现在检索结果中就满足检索条件, 这对检索者来说具有较大的积极意义, 他们可以通过逻辑“或”, 实现多方面的检索, 既可以得到想要的资料和数据, 也可以通过侧面的搜索来获得全面的信息, 提升了检索效率。第三种运算符是具有概念删除关系的一种组配, 逻辑“非”是两个检索词中应从第一概念中排除第二个概念, 通过这种检索方式, 能够在总体上实现一个较大的提升, 并且在检索的查

作者简介 欧丽清(1960-), 女, 馆员, 研究方向: 图书馆管理。  
收稿日期 2014-03-20

准率和查全率方面都取得一定的提升。

## 1.2 截词符

截词符也称统配符,用来对检索词(干)进行扩展,用“?”或者“\*”表示。截词符主要有前方截词、中间截词、后方截词,后方截词又分为后方无限截词以及后方有限截词。以前方截词为例,主要是将截词符放在一个词的前面,以表示其左边不管截去有限或者无限个字符,只要数据库中具有与截词符后面部门字符串相同的检索词的信息,即为命中信息。由此可见,截词符对文献检索具有非常重要的意义,在很大程度上能够实现全面的检索。

## 1.3 位置逻辑运算符

位置逻辑运算符是表示其连接的两个检索词之间的位置关系,常用的有(W), (nW), (N), (nN), (L), (S), (F)等。相对于前两种运算符来说,位置逻辑运算符具有很大的不同,而且能够避免模糊搜索带来的负面影响。位置逻辑运算符在实际的应用当中,类别较多,可以根据检索者的具体要求来进行检索,最大限度地提升了查准率与查全率。从现有的情况来看,在文献检索中应用位置逻辑运算符,首先可以提升查全率,为检索者提供更加广泛的信息范畴;其次,文献检索在应用位置逻辑运算符能够提升查准率。一般情况下,检索者在输入查找的信息以后,会得到很多没有用的信息,在大量的信息中寻找想要的数据和资料是非常困难的,位置逻辑运算符告别了这种情况,实现了有效检索,并且可以智能化的过滤一些没有用的信息。

# 2 提高查全率与查准率的方法

## 2.1 扩大检索范围,提高查全率

在文献检索的过程中,查全率会直接影响到检索者是否能够检索到自己想要的资料,进而对检索者的工作和生活产生影响。查全率并不是移位的扩大检索范围,而是要在相关的范围内进行扩大,避免重复的数据和资料出现,在给予更多选择的同时,保证资料的正确性和完整性。扩大检索范围,提高查全率可以从以下几个方面出发。

### 2.1.1 考虑同义词和近义词

在文献信息检索的时候,尽可能全和多的列举课题中的每个检索词中的相关词、同义词、近义词,并用逻辑“或”连接成检索式。一般情况下,检索者在检索文献的时候,常常将整个课题复制到检索框中,经常导致检索出来的资料很少,而且在时间上也不理

想。另外,当检索者直接复制课题的时候,检索出来的资料不一定是检索者想要的。在以后的检索当中,为了有效地提升查全率,可以应用逻辑“或”来提高查全率。比方说在查找“计算机辅助技术的应用”这一课题时,可以采用同义词和近义词的方式进行查找,例如“CAD技术”、“多媒体技术”、“计算机系统”等等,通过这些近义词和同义词的查找,可以缩小查找范围,并且提高查全率,获得更加符合需求的资料和数据。

### 2.1.2 选择较大检索范围的字段

在文献信息检索中,为了扩大检索范围,经常要选择较大检索范围的字段。通过这种方式,可以避免模糊搜索带来的“大海捞针”问题。例如,在查找“我国抗震设计的要点”时,有些人也发表过一些权威的文献,但是并没有用这个名字,而是用了“抗震设计的关键点”、“抗震设计的优化方案”、“抗震设计的未来趋势”等等,这些字段的检索范围较大,而且符合“抗震设计要点”的相关需求,同时能够帮助检索者获得更多的权威数据、资料。

## 2.2 缩小检索范围,提高查准率

在文献检索的过程中,为了有效的提高查准率,需要在一定程度上缩小检索范围,不需要的资料和数据会影响检索者的检索效率和质量。在缩小范围的措施当中,可以从以下几个方面出发。

### 2.2.1 使用逻辑“与”、逻辑“非”限制检索范围

使用逻辑“与”,可以缩小信息资源的检索范围,提高查准率。例如,检索式:“图书馆 AND 建筑”,检索到的既是“图书馆”方面又是“建筑”方面的文献。使用逻辑“非”,可用于排除不希望出现的检索词。从以上的表述来看,缩小检索范围能够对文献检索产生较大的积极意义。从客观的角度来说,逻辑“与”和“非”是对文献检索的一个有效过滤,能够直接帮助检索者扫除一些不必要的信息,在精准定位的情况下,实现有效检索;从主观的角度来说,检索者在查找文献的过程中,都希望一次性检索成功。所以,采用逻辑“与”、逻辑“非”能够有效的限制检索范围,并且满足检索者的主观需求。

### 2.2.2 使用位置逻辑运算符

相对于一般的运算符来说,使用位置逻辑运算符拥有自己独特的优势,采用位置逻辑运算符,能够对复合检索词进行加工修饰,限制词与词之间的位置关系,可以弥补布尔逻辑运算符只是定性规定检索词的

范围,使检索结果的查准率提高。文献检索的目的在于,通过合理的方式以及正确的关键词进行检索,从而得到想要的数据和资料。但是,大部分的检索者都不能准确的定义关键词,而且在前几种的检索方式中,都是从最开始的生疏到后来的熟练。在这个过程中,检索者会浪费大量的时间,而是用位置逻辑运算符可以帮助一些操作生疏的检索者在短时间内获得自己想要的资料。

### 2.2.3 使用下位词

除了上述的一些方法之外,检索者还可以通过下位词来缩小检索范围,进一步提升查准率。比方说,我们在检索“飞行器”这个词的时候,能够得到上千条记录,而检索“航天飞机”的时候,仅仅能够获得上百条的记录。从下位词的定义来说,“航天飞机”属于“飞行器”的下位词,检索者通过下位词进行检索,能够进一步确定自己想要查找的资料和数据范畴,无论是操作生疏还是操作熟练的人,都能够查到自己需要的数据和资料。从以上的表述来看,要想提高文献检索中的查全率和查准率,完全可以通过一些细节方面的改变而达到。未来的文献检索应该朝着更加智能化的方向发展,而不是一味地通过简单措施来弥补固有的缺陷。

## 3 对文献检索的思考

当下的文献检索集中在热门文献当中,而工作人员常常忽略冷门的文献或者一些关键词。在以后的工作当中,需要通过不同的措施和方式,全面提高文献检索的查全率和查准率。社会的发展速度在不断的加快,文献的数量和类别也会更加健全。单单针对热门文献提高查全率和查准率,并不能很好地帮助检索者来检索文献。另一方面,文献检索的相关信息和资料需要具备一定的密切关联,题目上的相同词汇并不意

味着内部信息的健全和完善,在以后的工作当中,提高文献内部的查全率和查准率,也是非常重要的。在日后的工作当中,必须通过多元化的检索方式,提升各项工作的水平和效率,达到一个理想的效果。

## 4 结束语

本文就如何提高文献检索中查全率和查准率进行了一定的探讨,从现有的情况来看,文献检索还有很多的问题没有解决,在后续工作当中,需要通过各项工作来实现一个较大的提升,满足检索者对文献检索的各项需求。相信在日后的工作当中,文献检索的查全率和查准率会有一个更大的提升,促进多项工作的进步,为检索者的工作和生活产生积极意义。

## 参考文献

- [1] 程琳.网络信息检索策略与检索途径探讨[J].图书情报论坛,2010,(3):49-50.
- [2] 余丹.关于查全率和查准率的新认识[J].西南民族大学学报(人文社科版),2009,(2):283-285.
- [3] 陈予琳.关键词检索方法在科技查新中的应用研究[J].河南师范大学学报(自然科学版),2011,(3):171-173.
- [4] 王小利,曹导叶,王杉.科技查新中查全率与查准率的辨析[J].科技情报开发与经济,2010,(14):139-140.
- [5] 李璐,江葆红,孙红红.搜索引擎检索策略调整研究[J].内蒙古科技与经济,2011,(14):81.
- [6] 黄尤来.建立全国查新报告数据库的作用及问题[J].河南科技,2010,(20):4,21.
- [7] 吴红莉,王文嘉.“南方谈话”相关文献分析——基于文献计量学[J].中共山西省委党校学报,2013,(1).
- [8] 赵洲.公共图书馆实行文献采购招投标利弊与对策探讨[J].晋图学刊,2013,(1).