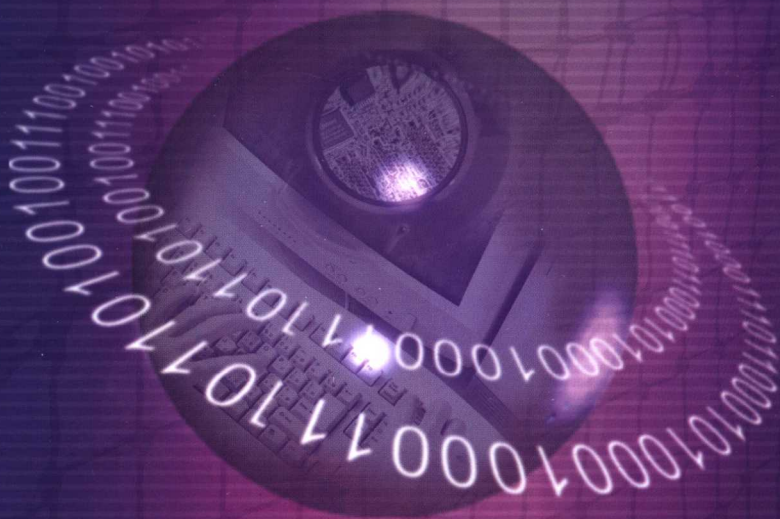


# 生物医学信息检索

SHENGWU YIXUE XINXI JIANSUO

王秀平 主编



 科学技术文献出版社

# 生物医学信息检索

主 编 王秀平

副主编 孔瑞珍 贺培风 韩玲革

编 者 (以姓氏笔画为序)

王 轶 冯 毅 宁红梅

孙 焱 乔建芳 李雪琴

苗艺青 武建光 袁永旭

科学技术文献出版社

Scientific and Technical Documents Publishing House

北 京

**图书在版编目(CIP)数据**

生物医学信息检索/王秀平主编.-北京:科学技术文献出版社,2004.7(重印)  
ISBN 7-5023-4516-7

I. 生… II. 王… III. 生物医学工程-情报检索-教材 IV. G252.7

中国版本图书馆 CIP 数据核字(2003)第 124748 号

出 版 者 科学技术文献出版社  
地 址 北京市复兴路 15 号(中央电视台西侧)/100038  
图书编务部电话 (010)68514027,(010)68537104(传真)  
图书发行部电话 (010)68514035(传真),(010)68514009  
邮 购 部 电 话 (010)68515381,(010)58882952  
网 址 <http://www.stdph.com>  
E-mail: stdph@istic.ac.cn  
策 划 编 辑 科 文  
责 任 编 辑 周 玲  
责 任 校 对 赵文珍  
责 任 出 版 王芳妮  
发 行 者 科学技术文献出版社发行 全国各地新华书店经销  
印 刷 者 富华印刷包装有限公司  
版 ( 印 ) 次 2004 年 7 月第 1 版第 2 次印刷  
开 本 787×1092 16 开  
字 数 335 千  
印 张 15  
印 数 8001~11000 册  
定 价 19.00 元

© 版权所有 违法必究

购买本社图书,凡字迹不清、缺页、倒页、脱页者,本社发行部负责调换。

# 序

新世纪伊始,人类大规模、高速度地进入信息和智能时代的步伐铿锵有声,传统的物质生产也正在让位于以知识密集、技术密集和智力密集为特征的知识产品的生产,一个崭新的知识经济时代正在到来。计算机技术、现代通讯技术和多媒体技术的高度发展和有机结合,借助知识爆炸所引发的信息革命洪流一浪高过一浪,正以空前未有的迅猛之势席卷全球,荡涤着世界旧有的一切积淀,冲击着人类社会的所有方面,迅速改变着人们的生产方式、生活方式、学习方式、工作方式、交往方式,乃至思维方式。

计算机网络飞速发展,现在几乎无所不在,数以千计的网络、数以万计的数据库和数以亿计的网页加上海量的网上资源为人们获取信息提供了极大的方便。为了快速、高效、准确地查找信息,信息检索技术和手段也在不断发展之中。在信息时代,是否掌握先进的信息检索技术,能否使用计算机通过网络查找所需要的知识无疑是判定一个科技人员基本信息素质的主要标准。谁掌握了信息检索技术,就等于掌握了信息;谁掌握了信息,就等于掌握了科技工作的主动权。

在 21 世纪的科技领域中,生命科学和信息科学都将是十分活跃的主体和前沿科学。生物医学领域的最新动态和研究成果等信息成为人们关注的焦点。如何快速准确地获取这些信息,是每个专业人员所面临的问题。《生物医学信息检索》就是为生物医学科技人员提供的一本简单易懂、实用性强的专业信息检索教科书。本书的作者有着丰富的从事生物医学信息检索和教学的实践经验,结合大量的最新资料,编写了这本符合医学高校本科生和研究生使用要求的教材,旨在培养学生和广大生物医学科技工作者的信息素养,提高大家利用现代信息检索工

具进行检索的意识和技能。该书对于网络信息检索技术的介绍新颖详细,很有特色。如网上常用免费数据库、网络特种文献、网上循证医学检索系统等内容的介绍,详细具体,是其他教科书所不及的。

中华医学会信息学会 主任委员  
卫生部情报图书专家咨询委员会 主任委员 **王汝宽**

# 前 言

医学信息检索课是培养学生情报意识、情报素养,学习和掌握现代化信息技术,提高开发和利用信息资源能力的一门科学方法课。随着现代科学技术的飞速发展,各种中外文检索数据库不断问世,尤其是 Internet 网络资源的大量涌现,使传统的医学文献检索课受到了前所未有的冲击,利用计算机进行信息查询已成趋势。因此计算机检索尤其是网络信息的检索已成为医学信息检索课的主要内容。

现代高新技术的发展使信息记录、信息传播和信息检索都发生了革命性的变化。近年来中国教育科研计算机网(Cernet)与国内各高校的连接,为高校教学、科研、医疗等提供了便利的通道和丰富的资源。校园网的建设和数字化图书馆的利用使广大师生员工的信息素质不断提高,利用计算机网络获取医学信息的积极性空前高涨。作为专业信息人员,看到这些感到十分欣慰,同时也认识到帮助广大师生员工掌握现代信息技术,更好地利用信息资源是我们义不容辞的责任。

为了满足研究生、本科生以及广大医务人员学习计算机信息检索以及医学信息知识的迫切需要,2001年我们编写了《医学信息检索与利用》教材,在不到一年的时间内即已告罄。我们总结了一年来的教学实践以及用户培训中的经验和信息反馈,针对当前师生最迫切的需要,重新编写了新的医学信息检索课教科书——《生物医学信息检索》。本书结合信息检索技术的最新发展,以全新的角度、新颖的内容阐述基本检索知识,反映学科新进展。与前《医学信息检索与利用》和其他医学信息检索教科书相比,具有以下特点:①内容新颖,增加了生物学和循证医学等内容,全部摘自最新的检索工具;②编排合理,从学生学习和使用的角度出发,采用符合学生学习规律的混合编排法,手工版和机检版相结合,容易掌握;③重点突出,对很多重要的内容如检索系统和数据库做重点介绍,如网络的免费数据库;④适用性强,本书不仅适用于研究生、本科生的医学信息检索课的教学,同

样适用于广大在职医务人员。可以满足他们学习基础检索知识和掌握计算机光盘和网络检索技术的要求。尤其是第九章“循证医学信息检索”,既是在校学生应当学习和掌握的重要知识,又是医务人员进行临床实践必需的内容。

参与编写本书的人员全部是长期从事医学信息检索教学或医学信息资源管理的专业人员。他们将多年来丰富的教学经验和工作中大量的实践积累融入教材的编写。经过大家的努力,新教材力求达到实用、简洁、与时俱进。本书共分为九章:第一章介绍了生物医学信息检索的基本理论和知识,包括机检基础知识;第二章、第三章介绍了中外文生物医学检索工具及国内外常用的计算机光盘数据库和全文数据库的使用;第四章介绍了 Internet 网络信息资源的基础知识,包括国内外著名医学网站、医学搜索引擎;第五章重点介绍了网上免费及常用生物医学数据库的使用;第六章介绍了网络出版物的使用和检索,包括免费生物医学电子期刊与电子图书的使用;第七章为特种文献检索,重点介绍会议文献、专利文献等的网络检索方法;第八章介绍目前大学图书馆的服务、职能、资源建设等,是学生更好地利用图书馆的必需知识;第九章介绍循证医学信息的检索,尤其是循证医学资源的类型、来源及使用方法。

非常感谢卫生部情报图书专家咨询委员会主任委员、中华医学会信息学会主任委员王汝宽教授在百忙中为本书作序。

本教材的编写和出版得到许多前辈和师长的指导、支持,同时也得到许多同事、朋友的无私帮助,在此一并表示衷心的感谢。

由于编者的知识与水平有限,虽想表达最新最好的内容,但终因篇幅有限且时间仓促,没有把全部想法和愿望在书中充分体现出来。编写中难免存在很多错误和不足之处,恳请同行和使用者批评指正。

编者

2003年11月

(京)新登字 130 号

## 内 容 简 介

本书是详细介绍生物医学信息检索的一本教材,内容主要包括生物医学信息检索的基本理论和知识、中外文生物医学检索的主要工具,Internet 网络信息资源的基础知识,网上免费及常用生物医学数据库的使用,网络出版物的使用和检索,特种文献检索,图书馆的利用和循证医学信息检索。适合广大研究生、本科生的医学信息检索课的使用,同样适用于广大在职医务人员。

---

科学技术文献出版社是国家科学技术部系统惟一一家中央级综合性科技出版机构,我们所有的努力都是为了使您增长知识和才干。



# 目 录

<b>第一章 信息与文献</b> .....	( 1 )
第一节 信息、知识、情报与文献 .....	( 1 )
第二节 文献概述 .....	( 3 )
第三节 信息检索基础 .....	( 5 )
第四节 信息检索工具 .....	(15)
第五节 计算机信息检索概述 .....	(17)
<b>第二章 中文医学信息检索工具</b> .....	(25)
第一节 《中文科技资料目录:医药卫生》.....	(25)
第二节 中国生物医学文献数据库(CBMdisc) .....	(28)
第三节 中文科技期刊全文数据库(VIP) .....	(33)
第四节 中国学术期刊全文数据库(CAJ) .....	(36)
第五节 《中国药学文摘》及其数据库 .....	(39)
第六节 其他中文检索工具 .....	(43)
<b>第三章 国外医学信息检索工具</b> .....	(48)
第一节 《IM》及 MEDLINE .....	(48)
第二节 《EM》及 EMBASE .....	(62)
第三节 《BA》及 BIOSIS Previews .....	(68)
第四节 美国《化学文摘》 .....	(78)
第五节 美国《科学引文索引》 .....	(88)
第六节 其他外文医学检索工具 .....	(93)
<b>第四章 网络信息资源检索</b> .....	(100)
第一节 网络信息检索基础知识.....	(100)
第二节 网上搜索引擎.....	(104)
第三节 生物医学网站.....	(114)
<b>第五章 网络医学检索系统及数据库</b> .....	(123)
第一节 PubMed 系统 .....	(123)
第二节 NLM Gateway 系统 .....	(131)
第三节 OVID 系统 .....	(134)
第四节 Entrez 系统.....	(140)

---

第五节	MicroMedex 医药信息系统 .....	(148)
第六章	网络电子出版物 .....	(154)
第一节	概述 .....	(154)
第二节	网络电子期刊 .....	(157)
第三节	网络电子图书 .....	(170)
第七章	特种文献检索 .....	(177)
第一节	会议文献检索 .....	(177)
第二节	专利文献检索 .....	(182)
第三节	学位论文检索 .....	(190)
第四节	技术标准检索 .....	(194)
第八章	图书馆的利用 .....	(197)
第一节	概述 .....	(197)
第二节	图书馆馆藏的利用 .....	(199)
第三节	电子图书馆 .....	(202)
第四节	参考工具书 .....	(204)
第九章	循证医学信息资源的检索 .....	(212)
第一节	概述 .....	(212)
第二节	循证医学信息资源类型及检索特点 .....	(214)
第三节	循证医学检索系统及数据库 .....	(216)
参考文献	.....	(228)

# 第一章 信息与文献

## 第一节 信息、知识、情报与文献

### 一、信息

信息是事物存在和运动状态及其特征的反映,即客观事物所发出的信号和消息。信息普遍存在于自然界、人类社会以及人类的思维活动中。事物在不断变化,信息也在不断产生。不同事物具有不同的存在方式与运动状态,形态特征各不相同,因而会发出不同的信息。人们就是通过获取、识别自然界和社会的不同信息来区别事物,从而认识世界和改造世界。根据来源不同,人们一般将信息分成四种类型:自然信息、生物信息、社会信息和机器信息。如风云雨雪、湖光山色是自然信息,花红柳绿、血压升降是生物信息,人们使用的文字、语言、图形符号是社会信息,机器发出的脉冲信号等是机器信息。

随着自然界和人类社会的不断发展,信息的产生和利用达到了空前的高潮,因而有人将目前的时代称为“信息时代”。信息已和物质、能源一起构成了人类社会的三大支柱,信息在人们日常生活、社会活动、科研活动中的作用越来越大,甚至将一个国家信息化程度的高低作为衡量这个国家发达水平的重要标志。

### 二、知识

经过人脑的选择整理,系统化、理论化的信息就是知识。知识是人类在认识世界和改造世界的过程中所获得的认识和经验的总和,是人的主观世界对客观世界的如实反映和概括。人们在改造自然和社会的实践过程中,通过信息不断感触和认识世界,掌握自然界、人类社会和人类思维方式的发生、发展规律,并对感觉到的客观事物进行总结、加工、提炼、系统化、理论化,从而形成知识。也就是说,知识是人的大脑通过思维重新组合的系统化的信息集合,是信息的一部分,是经过优化以后的理性化的信息。

一般来说,人们把知识分为自然科学知识、社会科学知识和哲学知识三大类。

### 三、情报与情报意识

“情报”一词,最早源于军事,即战时敌方的情况报道,多带有机密的性质。随着社会和科技的不断发展,情报的概念也发生了变化。目前多数学者认为,情报是指针对一定的需要通过

一定方式传递而获得的能解决一定问题的有用知识,是活化的,可产生一定效益的知识。这种知识可以是别人传达的,也可以是自己通过调查、学习或接受教育得来的。现在,情报已经广泛渗透到经济、政治、科技、市场、商品等领域,成为人们解决某一特定问题所需知识的代名词。

情报具有三个基本属性:一是知识性,即情报的本质是知识,知识是情报的实体;二是传递性,知识通过交流和传递,得到活化后才能成为情报;三是效用性,即情报的应用都应产生一定的效益,是情报使用价值的体现。也就是说,沉睡的知识不是情报,只有被激化、活化、利用的知识才成为情报。近来竞争性作为情报的第四个属性已被人们不断的提及和认可。

情报意识即“意识情报的能力”,是人们对大量客观存在的情报的自觉反应。在人们的日常生活、工作、科研活动中,有大量亟待解决的问题、难题,客观上也存在大量的能解决这些难题的知识。能够感知这些知识的存在,并将它们运用于实际,解决相应的问题,就是情报意识在起作用。同时,情报意识也指人们获取情报的积极性和主动性。每当工作、学习中遇到难题时,能积极主动地查找、获取情报,也是情报意识的具体体现。一个人情报意识的强弱,与他的工作责任心、对待工作的态度等有极大关系。工作责任心强的人,会时时留心、处处在意与工作、科研息息相关的知识,从而获得较多解决问题的情报,情报意识就强;反之,对工作采取无所谓、得过且过态度的人,即使情报就在眼前也会视而不见、听而不闻,情报意识就弱。我们学习信息检索这门课程,一方面要掌握获取情报的途径与方法,更重要的是要培养学生的情报意识,树立积极的人生观和工作态度,为社会进步和科学发展做出更大的贡献。

## 四、文献

利用不同的手段将知识记录到各种载体上就形成文献。换言之,记录知识的一切载体就是文献。

文献是人类长期从事生产和科学研究活动、社会活动的真实记录,凝结着亿万人的劳动和智慧,记载了许多有价值的事实、数据、理论、方法、成功的经验、失败的教训,反映了人类历史上各个时期科技发展和社会活动所达到的成就和水平,预示着将来的发展趋势和方向。我们工作和科研所需的大量信息就记载在文献上,它是一种重要的信息资源、科技资源、精神财富。

文献由三方面要素构成:知识、载体和记录的手段,三者缺一不可。知识是文献的主体和灵魂,载体是文献的具体形式,记录手段是联系知识和载体的桥梁。由此可见,文献和知识既是不同的概念,又有密切的联系。文献必须包含知识内容,而知识内容只有记录在物质载体上,才能构成文献。因而,人们对信息和知识的需求常常转化成对文献的需求,在大多数情况下,人们从获得的文献中获取对自己有用的信息。

## 五、信息、知识、情报、文献间的关系

从信息、知识、情报和文献的概念来看,它们之间有着极为密切的关系。信息是事物存在和运动状态的反映,它无处不在,无时不在;知识是人类在实践中对大量信息进行加工、提炼、系统化后形成的,是优化的、理性化的信息;情报是被激活的、能产生效益的知识,是知识价值的具体体现;文献是记录在不同载体上的知识,是信息、知识、情报存储的主要方式。信息包含

着知识,知识包含着情报,文献则是信息、知识和情报的外在表现形式。

(孔瑞珍)

## 第二节 文献概述

### 一、现代医学文献的主要特点

随着生产力水平的不断发展,科学研究出现了既高度分化,又高度综合的局面。学科分支逐渐增多,且各分支之间相互交叉渗透,造成边缘学科和新兴学科不断涌现。随着科学技术的迅猛发展,知识的门类和数量日益增多,用来记载和传播知识的文献随之出现了极其复杂的局面。就生物医学文献而言,主要有以下几方面特点:①数量庞大,增长速度加快,居各学科之首。②内容交叉重复。随着科学技术的不断发展,学科间综合交叉,彼此渗透,同一文献可用不同形式、不同文字在不同范围内多次发表。③分布分散。当前医学学科越分越细,内容渗透交叉,加上文献出版的种类越来越多,造成各专业学科的文异常分散。文献分布的分散性导致了信息检索全面性的下降。④文种繁多。据报道,目前世界出版文献的文种已达70~80种之多。⑤出版类型多样化。科技的发展,使记载知识的载体多样化,进而使文献类型日趋复杂。目前,除传统的印刷型文献外,视听型、缩微型、机读型文献相继问世。⑥文献老化速度加快。由于科学技术发展迅速,科技论文数量急剧增加,旧的知识、理论、方法更快地被新的内容所取代,科技论文也随之很快失去时效。

综上所述,当今的医学文献呈现出数量庞大、复杂多样的局面。正是由于文献这种特点,检索成为人们获取文献信息的必经之路,成为人们学习、科研、创新的有力工具。

### 二、文献的分类与分级

#### (一)文献的分类

按照载体不同,可对文献进行如下分类:

##### 1. 印刷型文献

以纸张为载体,通过印刷的方式将知识固化在载体上的文献。这是一种传统的文献类型,因符合人们的阅读习惯而成为目前使用最多的文献形式。我们常见的图书、报刊等都是印刷型文献。但这种文献的存储密度低,而且易燃,害怕虫蛀、鼠咬、水蚀,不易保存。印刷型文献按其出版方式不同,又可分为以下三种:

##### (1)图书

图书是将生产技术、科技成果进行总结和对人们社会活动进行概括的文献,系统论述某一个方面的专题,内容比较成熟、系统、稳定,但一般不能反映最尖端的科技成果和社会发展水平,是印刷型出版物中最普遍的一种文献。根据内容与作用图书又可分为一般性图书(如专著、文集、教科书、丛书等)和参考工具书(如词典、百科全书、指南、手册)两大类。

## (2) 期刊

是定期或不定期的连续出版物,具有固定的名称、编辑机构、装帧版式和结构编排,内容专深新颖、信息量大、报道迅速,能反映当前的科技水平和社会发展状况,是最主要的信息源,也是信息检索的主要对象。

## (3) 特种文献

又称非书非刊出版物,是除图书、期刊以外的其他出版类型的总称,有科技报告、专利文献、会议文献、学位论文、政府出版物、标准文献、技术档案、产品资料等类型。这类文献的特点是一般不公开出版,普通图书馆也不收藏,但特种文献反映了许多最新科学研究和技术以及国家的法规、标准等不可或缺的信息,也是重要的和新颖的情报源。

### 2. 缩微型文献

这是以感光材料如胶片、胶卷为载体,将印刷型文献通过摄影技术缩小后形成的文献。由于目前摄影技术发达,小小一张胶片上可以存放成千上万张印刷型文献,因此,这类文献的存储密度高于印刷型文献,且易复制、收藏,但缩微型文献不方便阅读,必须借助特定设备——缩微阅读机才能阅读,流传并不广泛。

### 3. 视听型文献

又称声像型资料,是以磁性材料如磁带、录像带等为载体,将文献以声音、图像等方式录制而成的文献,包括唱片、录音带、电影片、幻灯片等。这是一类用途特殊的文献,可闻其声,观其形,具有形象、直观的特点,在特殊的场合(如教学实践、临床观摩等)具有独特的作用。

### 4. 机读型文献

即计算机可以阅读的文献。载体为计算机存储介质,如计算机硬盘、软盘、光盘等。这种文献主要采用磁性存储技术,通过编码程序设计,把文献内容变成逻辑语言和机器语言,输入到计算机中进行存储,阅读时再由计算机进行输出。机读文献根据文献形式可分为数据库文献和网络文献;根据构成的语言种类又可分为文本型文献、多媒体型文献、超文本型文献和流媒体型文献,其中超文本型文献是网络传输的专用文献类型,也是 Internet 可以识别的文献类型,流媒体文献则是图像和音频在线传输和实时接收的专用文献。机读文献存储密度高,通过计算机程序可随意处理和使用,并可通过网络进行传输,共享性好,有独特的优点。而且随着计算机和 Internet 的普及,这种类型文献的使用必将日趋广泛。

## (二) 文献的级别

按照文献加工程度的不同,可对文献进行如下分级:

### 1. 一次文献

一次文献指未经加工的原始文献,是作者以生产或科研成果为依据创作而成的文献,是创作者对其新发明、新创造、新技术、新方法、新见解等的真实记录,比较具体、详尽、系统,它的特点是具有创新性。专著、期刊论文、专利文献、技术标准、科技报告、学位论文、会议论文等都属于一次文献。

### 2. 二次文献

二次文献是对分散的无组织的一次文献进行分析、加工、提炼、排序整理后形成的文献。

一般将原始文献提炼出主要的外表和内容特征,然后排序而成。它是为了便于管理和利用一次文献,由情报人员编辑、出版和积累起来的工具性文献。二次文献包括目录、索引(题录型检索工具)、文摘(文摘型检索工具)。由于二次文献中包含大量内容相关的一次文献,因而成为信息检索的主要工具。

### 3. 三次文献

三次文献是为了一定的目的,在利用二次文献的基础上选用大量与某专题相关的一次文献,将其内容进行分析、整理、综合后形成的文献。三次文献是对某一个学科或专题的概括性叙述,常见的有专题综述或述评、年鉴、手册、进展、指南等,是人们从总体上把握科技动态的重要工具。

从一次文献到二次文献、三次文献,是一个由分散到集中,由无组织到系统化的过程,对广大科技人员和图书情报工作者来说,一次文献是主要检索对象,二次文献是进行信息检索的主要工具,三次文献则是情报调研的结果,是把握科技动态的有力工具。

另外,尚有零次文献的提法,指尚未形成文字的材料或知识,即非出版型或口头交流的文献。这是一种特殊类型的文献,也是非常重要的情报源。但目前图书情报界对此尚有争议,在此不作详细论述。

(孔瑞珍)

## 第三节 信息检索基础

### 一、信息检索的概念和作用

#### (一)信息检索的概念

信息检索的概念有广义和狭义之分。广义的信息检索包括信息的存储和检索,是指将信息按一定方式组织和存储起来,并根据用户需要查找到有关信息的过程和技术。信息存储的过程包括信息的收集、标引、著录,由情报工作者去完成。狭义的信息检索仅指检索,即从大量的信息中迅速、准确、全面地查找出特定信息的过程。通常所说的信息检索一般是针对其狭义概念而言。

信息检索按检索对象不同,可分为数据检索、事实检索和文献检索。数据检索是以特定的数值型数据为检索对象,它包括各种统计数字、图表、化学结构式、计算公式等等,如某种药物的理化常数、常用剂量、结构式等;事实检索是以特定的事实为检索对象,如什么是艾滋病,我国首例艾滋病是何年发现的等等。文献检索则是从大量的文献中将涉及某一特定课题的文献查找出来。虽然文献检索仅是信息检索的组成部分,但在整个信息检索中,文献检索是主要的。人们对数据和事实的检索,在很多情况下都要借助文献检索,因为目前文献仍是存储信息、数据、事实的主要形式。

## (二)信息检索的作用

掌握信息检索的方法与技能,可以有效借鉴前人的科研成果和经验,能够增强获取所需信息的意识。科学发展的每一步都依赖和借鉴前人的成果、经验和教训,而借鉴和交流最主要是靠文献的记载和信息的传递。文献和信息浩如烟海,要迅速、准确地找到科研必需的内容,信息检索是捷径。所以说,信息检索是联系情报用户和信息的桥梁,是打开世界信息宝库的金钥匙,能为科研工作提供参考和决策。

另外,经常不断地进行信息检索,可以化知识为能力,为改善智能结构提供有利条件。智能是指人类认识客观事物、运用知识解决实际问题的能力,具体来说,指自学能力、研究能力、思维能力、表达能力等等。这些能力是人们成长、成才、成就事业所必需的。

总之,通过信息检索的学习与实践,不仅掌握了必要的知识,训练了方法技能,而且可以找到求知的钥匙,增强情报意识,改善智能结构,收到事半功倍的效果。

## 二、信息检索的一般原理

目前,信息检索有手工检索和计算机检索两种方式。手工检索,就是利用印刷型检索工具进行文献检索的方式;计算机检索则是利用计算机对存储于数据库或网络中的文献进行检索的方式。两种方式的检索原理基本一致。

信息检索的全过程包括文献信息的存储和检索两个过程。存储过程,主要是对文献信息进行标引,形成文献信息的特征标识,为信息检索提供有章可循的途径的过程。文献信息的特征标识,有代表文献外表特征的如标题、作者、来源、语种等项目,也有代表文献内容特征的如文献论述的主题、分类等项目。然后将此特征标识利用特定检索语言转化成检索工具使用的特定标识——存储标识,最后按照存储标识的顺序进行编排。存储结果的表现方式是印刷型检索工具如目录、索引、文摘或计算机信息检索工具——数据库。检索过程是根据信息需求,将检索课题转化成检索提问词或提问式,如著者姓名、主题词、分类号等,并将提问词通过特定检索语言转化成某种检索工具可以使用的检索标识,按此标识到检索系统中查找文献信息的线索,并索取原始材料。文献信息存储和检索是两个相互匹配而且互逆的过程,存储是检索的基础,检索是存储的目的,二者缺一不可。在这两个过程中,用户使用的检索标识必须与检索工具使用的存储标识相一致才能查到符合课题的文献信息。在此,检索语言起到了桥梁和纽带的作用。

## 三、检索语言

从信息检索的原理可以看到,用户要检索到适当的文献信息,关键是要使自己使用的提问标识与检索工具使用的检索标识一致。二者达到统一的关键在于使用相同的检索语言。因此,检索语言在检索过程中的作用非常重要。

所谓检索语言,是指在信息检索过程中用来描述文献特征和表达情报提问的一种专门的人工语言。它是掌握文献信息检索技能所必须具备的知识,是统一文献信息标引人员和检索者使用检索用语的桥梁。



检索语言根据是否受过人为规范可分为受控语言(规范化检索语言)和自然语言(非规范化检索语言)两大类。受控语言是指检索用语经过人为控制和规范的检索语言,是把检索用语中的同义词、近义词、同物异名词等进行规范化处理,使每个检索词只能表示一个概念,每个检索概念只用一个惟一的检索词来表达,从而提高信息检索效率。检索中经常使用的主题语言、分类语言等都是受控语言。自然语言与受控语言相反,它对检索用语中的同义词、近义词等不作处理,而直接使用原始语言作为检索语言。检索中常用的关键词语言、著者语言等是常见的自然语言。

检索语言根据构成原理可分为三类:分类语言、主题语言和代码语言。

### (一)分类语言

分类语言是按照检索用语在某种分类系统中的学科位置和属性来确定的,一般又分为体系分类语言和组配分类语言两种。

体系分类语言是体系分类法中使用的检索语言,是以学科专业为基础,根据文献的内部和某些外部特性,运用概念划分的原理,将知识门类按照从一般到具体、从简单到复杂、从低级到高级的原则分门别类地排列起来,层层划分,逐级展开,用等级来表示类目的从属关系、列举来表示类列的完整性的,因而又称为等级体系分类法。体系分类能较好地体现学科的系统性,揭示知识之间的隶属、派生和平行关系,便于从学科专业的角度集中、整理和查找文献信息。这种分类法的典型代表是国际著名的杜威分类法。但体系分类法不能很好地反映信息内容的多样性和交叉性,不够灵活方便。

组配分类语言是组配分类法中使用的检索语言。它是将信息内容所反映的客观事物分解成许多因素,再将这些因素归纳为若干范畴;对信息进行分类时,根据其内容的因素,利用范畴表中的有关类目,组配成适合特定内容的新类目。如阮冈纳赞编制的《冒号分类法》是组配分类法的典型代表。这种分类法能正确反映图书内容的多样性和交叉性,但系统性较差,不太适合于文献信息资料的组织编排。目前,组配分类法的优点正被体系分类法所吸收,如我国目前通用的《中国图书馆分类法》,以体系作为分类主导,但其中允许使用组配复分的方法集中某些类目。

### (二)主题语言

主题语言是依据检索用语的词义范畴、学科属性等,用词汇来表达概念的检索语言,它具有直观性、专指性、适应性、集中性、多元性的特点。主题语言按检索标识的来源不同又可分为标题词语言、单元词语言、关键词语言、叙词语言等。

标题词语言是指由事物名称作为主题而设立的检索语言。

关键词语言是指从文献信息中提取出来的,能表达文献信息的实质内容的名词术语,它能准确反映文献信息的内容实质,但由于没有经过规范,用户在检索文献信息时常常造成查找文献信息全面性的不足。

叙词通常称为主题词,因此叙词语言又可以叫做主题词语言,指能表达文献信息的实质内容,经过规范化处理的名词术语,由于经过规范化,所以查找文献的全面性比较强。另外还有