

含光盘

21世纪全国高等医药院校教材

供医药学各专业使用

医学文献检索

主 编 刘艳杰

中国医药科技出版社



21 世纪全国高等医药院校教材

医学文献检索

主 编 刘艳杰

编 委 (排名不分先后)

韩晓英 曹 荣 康 蕾 徐 宁

夏一鑫 秦 明 赵艳宏 范晨荟

郑 义 张爱华 朱 佳 刘光翀

中国医药科技出版社

·北京·

图书在版编目(CIP)数据

医学文献检索/刘艳杰主编. - 北京:中国医药科技出版社,2006.7

ISBN 7-5067-3480-X

I. 医... II. 刘... III. 医学-情报检索
IV. G252.7

中国版本图书馆 CIP 数据核字(2006)第 087802 号

出版 中国医药科技出版社
地址 北京市海淀区文慧园北路甲 22 号
邮编 100088
电话 010-62244206
网址 www.mpsky.com.cn
规格 787×1092mm 1/16
印张 14
字数 358 千字
版次 2006 年 8 月第 1 版
印次 2006 年 8 月第 1 次印刷
印刷 广州天河粤达印刷厂
经销 全国各地新华书店
书号 ISBN 7-5067-3480-X/G·0515
定价 28.00 元
本社图书如存在印装质量问题请与本社联系调换

出版说明

随着我国高等教育改革的深入，我国的高等医学教育在教学体制、教学理念、学科设置和教学内容等多方面都取得了长足的进步。21 世纪的医学教育将更加注重人才的综合培养：不仅要培养学生具有学科专业知识和能力，而且要具有知识面宽、能力强、素质高的特点，注重创新精神、创新意识、创新能力的培养。

教材建设是教学改革的关键环节。长期以来，医学教育教材的单一已不能体现各高校的办学特点，也不能体现教学改革与教学内容的更新。教材的多元化和具有地方性特色是教材建设的必要手段。因此，为了适应 21 世纪医学教育发展的需要，我们组织有关专家编写了这套“21 世纪全国高等医药院校教材”。

本套教材的编写是在充分向各医学院校调研、总结归纳的基础上开展的。在编写过程中特别注重体现各学科的基本理论、基本方法和基本技能，力求体现内容的科学性、系统性、实用性和可读性，最大程度地满足师生们的要求。在基本理论和基本知识上以“必须，够用”为度，并作适当扩展；重点强调基本技能的培养，突出实用性。本套教材紧扣人才培养目标和教学大纲，适当兼顾各校学生不同起点的要求，以确保教材的实用性和通用性，可供高等医药院校临床、基础、预防、护理、口腔、药学、检验、卫生管理等专业使用。

该套教材汇集了各学科相关专家多年来教学经验和实践经验，在编写过程中付出了大量心血，也做了很多有益的尝试和创新。衷心希望这套教材能够为我国的医学教育贡献一份力量。当然，由于时间仓促，不可避免地还会存在各方面不足，欢迎各院校师生批评指正。

编 者

2006 年 8 月

目 录

第一章 文献检索概论	1
第一节 文献检索概述	1
一、信息、知识、情报	1
二、文献和文献检索	3
三、医学文献的特点	3
四、文献检索的目的和意义	4
第二节 文献检索基础知识	5
一、文献检索的概念、原理和意义	5
二、检索语言	7
三、检索工具	9
四、检索方法、途径和步骤	12
第三节 文献的分类	15
一、按载体形式划分	15
二、按加工深度等划分	18
第四节 计算机文献检索	19
一、计算机检索的发展历史	19
二、计算机检索的特点	20
三、计算机检索系统	21
四、计算机检索方式	22
五、计算机检索的基本技术	23
第二章 国内医学文献检索	25
第一节 常用中文检索刊物	25
一、国内检索刊物的历史与现状	25
二、常用国内医学文献检索刊物简介	26
三、《中文科技资料目录》(医药卫生)	28
四、《国外科技资料目录》(医药卫生)	33
五、《中国医学文摘》	34
六、《全国报刊索引》	37

七、《中国药学文摘》	38
第二节 中国生物医学文献数据库	40
一、概况	40
二、数据库的检索运算符	41
三、CBMDisc 检索系统	42
四、数据库的检索方式	43
五、检索结果的输出	50
第三节 VIP 中文科技期刊数据库 (web 版)	52
一、概述	52
二、检索方法	52
三、输出方式	54
第四节 中国期刊全文数据库	54
一、概况	54
二、检索途径	55
三、检索结果输出	57
第三章 国外医学文献检索	59
第一节 美国《医学索引》	59
一、著录格式	59
二、检索途径	62
第二节 PubMed 检索	63
一、PubMed 数据库系统	63
二、检索方式	65
三、结果处理	69
第三节 WinSPIRS 系统的 MEDLINE	72
一、概述	72
二、数据结构	72
三、检索前的准备和主界面	74
四、检索方法	75
五、结果输出	79
第四节 OVID 系统的 MEDLINE	81
一、OVID - MEDLINE 数据结构	81
二、检索前的准备	84
三、检索方法和途径	85
四、结果处理	89
第五节 EBSCO - MEDLINE	90
一、检索前的准备	90
二、EBSCO - MEDLINE 检索	91
三、结果处理	94

第六节 美国《生物学文摘》与 OVID - BIOSIS	95
一、概况	95
二、期索引编排结构	95
三、OVID - BIOSIS	99
第七节 美国《化学文摘》	105
一、概况	105
二、期索引编排结构	107
三、文摘著录格式	111
四、期索引与检索方法	112
五、卷索引与检索方法	115
六、累积索引	118
七、指导性索引	118
第八节 其他英文医学检索数据库简介	119
一、Elsevier Science SDOS 全文数据库	119
二、SpringerLink	120
三、Kluwer Online Journals	121
四、Wiley Inter Science	122
五、ProQuest Medical Library	123
六、OCLC FirstSearch 检索系统	124
第四章 网络医学信息资源	126
第一节 网络信息资源概述	126
一、网络信息资源简介	126
二、网络信息资源分类	127
三、网上生物医学信息资源的分布	129
四、网络信息资源的特点	130
第二节 通用及医学专业搜索引擎	131
一、搜索引擎概述	131
二、医学专业搜索引擎介绍	134
三、综合搜索引擎	140
第三节 主要医学网站介绍	143
一、基础医学主要网站	143
二、临床医学主要网站	150
第四节 医学软件	158
一、医学软件网站	158
二、医学软件	159
第五章 引文索引	162
第一节 引文索引概述	162

第二节 中国科学引文数据库·····	162
一、概况·····	162
二、结构与检索方法·····	163
第三节 美国《科学引文索引》·····	165
一、《科学引文索引》简介·····	165
二、《科学引文索引》的作用·····	166
三、SCI 的结构编排·····	167
四、SCI 机检系统·····	172
第六章 特种医学文献检索·····	175
第一节 专利文献检索·····	175
一、专利·····	175
二、专利法和知识产权·····	176
三、国际专利分类法·····	178
四、INID 码·····	180
五、国内专利文献检索·····	180
六、德温特专利及其检索方法·····	184
第二节 医学会议文献检索·····	188
一、概述·····	188
二、国内会议信息检索·····	189
三、国外会议文献检索·····	192
第三节 学位论文检索·····	194
一、国内学位论文检索·····	194
二、国外学位论文检索·····	195
第四节 药物信息检索·····	196
一、常用药文学文献检索数据库·····	196
二、国外主要药学专业网站·····	197
三、国内药学网站·····	202
第七章 医学论文写作基础·····	204
第一节 医学论文的类型与特点·····	204
一、医学论文类型·····	204
二、医学论文的基本特点·····	206
第二节 医学论文的格式与内容·····	207
第三节 医学论文撰写方法及常见问题·····	212
一、撰写方法·····	213
二、写作步骤·····	213
三、写作的规范化问题·····	213

第一章 文献检索概论

【目标解读】

1. 了解信息、知识、情报的基本概念和关系
2. 理解文献的基本概念和特点
3. 理解文献检索基础知识
4. 了解文献的分类
5. 了解计算机文献检索的基本知识

第一节 文献检索概述

一、信息、知识、情报

在学习文献检索课的基本原理和检索方法之前，首先介绍文献检索的一些基本概念：信息、知识、情报以及三者的相互关系。

（一）信息

在我们所处的世界中，信息无处不在，无时不有。比如去股票市场可以了解股市信息；到人才市场可以了解人才需求信息；到电器商行可以了解电器信息；到软件市场可以了解软件信息；浏览经济导报可以了解经济信息等。信息不仅包括人与人之间的信号的交换，而且还包括人与自然界之间，人与机器之间，机器与机器之间的信号的交换，以及动物界与植物界的信号交换。由一个机体传递给另一个机体，甚至由一个细胞传递给另一个细胞，由一个分子传递给另一个分子的任何运动方式，也都被视为信息的传递。信息几乎存在于一切领域中。

随着信息时代的到来，更加丰富了信息学的理论。信息的内容逐渐增多，信息的形式可以图、文、声三种并存，信息的传递速度有了迅猛的发展。人们不仅从宏观的角度来研究信息，而且也从微观的角度来研究信息，大到天体，小到分子水平来研究自然界信息的规律和特性。

究竟什么是信息？不同学者从不同角度对信息做出各种定义，难以有统一的定论。哲学家认为信息是认识世界的依据；数学家认为信息是一种概率；物理学家说信息是“熵”；

而通信家则把信息看做是“不定度”的描述。信息论的创始人香农 (C. E. Shannon) 认为: 信息是组织程度, 它能使系统的有序性增强, 减少破坏、混乱和噪声。

信息是无形的, 它是客观事物的运动状态和特征的反映, 是人们认识事物发展的基础。每件事物不同的运动状态和特征会产生不同的信息反映, 包括社会信息、生物信息、技术信息等。在医学上, 各种体征的出现和变化是机体生理和病理状态的信息反映, 各种化验结果的数据也是机体生理和病理状态的信息反映。事物是不断变化发展的, 那么它的运动状态和特征也应随着不断变化而发展。因此, 信息也是不断变化发展的。但人们发现和认识信息都受着生产力、科技水平及认识手段的限制, 社会历史的发展是在不断获取信息、认识信息、传递和利用信息、创造信息的过程中, 通过信息认识世界与改造世界。在医学领域中通过各种疾病的不同信息来认识千差万别的疾病。

在人类已经进入信息社会的今天, 信息被视为如同能源、空气和水一样重要。信息资源被视为与自然资源同等重要的第二资源。人们将一个国家信息化程度的高低作为衡量一个国家发展水平的标志。一个国家科学技术发展水平、经济持续发展都与信息的开发和利用密切相关。

(二) 知识

汉语词典中, 知识解释为: 人们在改造世界的实践中所获得的认识和经验的总和。知识是人类对自然界、人类社会中各种现象、规律的信息反映进行思维分析、加工提炼、经过系统化、理论化的过程。也就是人的大脑通过思维重新组合的系统化的信息集合。因此, 系统化、理论化的信息就称为知识。信息是知识的源泉、材料, 知识是信息的理论和结果。例如某一症状、体征诊断为某一疾病, 这种症状和体征是该疾病信息的反映, 该疾病是症状和体征的信息升华, 这种信息升华就是疾病的诊断知识。因此, 信息是知识的源泉和基础, 知识是信息的升华。

(三) 情报

情报最初产生于军事领域, 主要是探察敌情的报告, 称之为情报。情报是人们在一定的时间内为一定的目的而传递有使用价值的知识或信息。这种情报具有保密性、时效性、传递性的特点。因此, 情报是知识的一部分, 是特定的知识或信息。随着历史、社会以及科学技术的不断发展, 情报的作用也不断地变化, 由军事转移到科技、经济、社会服务等各领域。现已转化为人们获取知识、信息的一种重要手段。

情报的种类有多种。按服务对象不同, 可分为军事情报、科技情报、战略情报、战术情报等; 按传递媒介分为文字情报、实物情报、声像情报; 按传递范围分为大众情报和专门情报。

(四) 信息、知识、情报三者的相互关系

信息是知识的源泉; 知识是系统化、理论化的信息; 情报是活化的知识和信息, 能为

人们所利用,其涵义关系是信息>知识>情报。信息要成为情报,一般要经过选择、综合、分析和研究加工过程,即经过知识的阶段才能成为情报,因此它必须具有三种基本要素:知识、传递和效益,知识是情报的实体,传递是情报的表现形式,效益是情报的结果。情报就是信息经过综合、筛选、逻辑思维、重新组合的系统化知识。

二、文献和文献检索

文献是记录有知识的一切载体的统称,即用文字、图像、符号、声音、音频、视频、编码等手段记录的人类知识的各种载体。

医学文献检索是一门以培养医学生的情报意识、掌握用手工或计算机方式从文献中获取知识和科技情报信息的工具性课程,是培养学生独立学习能力的一种专门知识。其目的是传授如何在知识的汪洋大海中获取所需目标知识,即学会查找文献,最终使文献为我所用,也是继续学习、提高自我的有效途径。

三、医学文献的特点

1. 数量庞大

医学文献是整个科技文献的重要组成部分,从整个世界科技出版物来看,医学文献占整个科技文献的四分之一。美国国立医学图书馆是世界上最大的生物医学文献信息中心,它研制开发的 Medline 数据库收录了来自 70 多个国家的 3700 种医学期刊,目前已有 1500 多万条文献记录,基本上囊括了世界上重要的医学文献信息。在我国目前有科技期刊 4400 余种,医学及相关学科的期刊就有 1200 种左右。中国生物医学文献光盘数据(CBMDisc)收录了医学及相关学科期刊 1000 多种,囊括了国内医学信息。医学文献数据库层出不穷,而且收录文献不断增加,医学文献增长十分迅速,数量非常庞大,占各学科之首。

2. 载体多样化

目前医学文献的载体除了传统的印刷型外,还有录像带、录音带、电影、幻灯等各种视听型;缩微胶片、缩微胶卷、缩微卡等缩微型;磁带、磁盘、光盘等电子型载体。载体的多样化,大大方便了读者。在相当长的时期内,印刷型文献将与其他类型文献同时并存,互相补充。

3. 语言障碍增加

随着科学的发展,医学文献出版的语言种类急剧增多,如美国《化学文摘》收录的语种有 50 多种,美国《医学索引》收录的语种有 44 种。在医学文献中英语出版物约占 70%。文种的增加丰富了医学文献,拓宽了研究空间,也造成了读者阅读文献的各种障碍,阻碍了科技情报信息的交流。

4. 重复发表,学科交叉与出版分散

由于受多种因素的影响,文献重复发表的现象越来越多,既是现代科学技术综合交叉、彼此渗透的反映,也有同一篇文章用不同形式、不同文字、在不同范围内多次发表的

现象。会议论文、学位论文、科技报告除了以自己独特的形式出版外,经过一段时间后又转化为其它形式出版。美国国家科学基金会的技术报告有 95% 在期刊上发表。11 种美国出版的检索工具引用的 17000 多种期刊中的重复率高达 50%, 美国《化学文摘》与美国《生物学文摘》重复率也高达 25%。这种重复大大增加了文献量的冗余,也会造成检索上的困难,浪费人力物力。各学科的交叉渗透,致使文献显得分散。一种专业刊物上报导的内容往往包含了多个学科,一篇专题论文又往往涉及几个专业,这是学科交叉渗透的必然结果。

5. 知识信息更新加快

科学技术的迅速发展,知识更新与信息交流速度加快,导致文献使用寿命缩短。文献的出版更落后于科学技术的发展步伐,有些文献还未出版或刚出版就被更新的知识所更替,新材料、新理论、新工艺、新方法不断推出,科技文献的老化周期已由原来的 50 年缩短到 5~10 年。由于在教学、科研和医疗过程中医学文献代谢频繁,大学生毕业后 10 年在学校里学的知识 80% 以上已陈旧过时,所以当今的每位医学工作者必须以终身学习取代一次性学习。

6. 时滞问题严重

科学论文数量的增加和更多论文被发表,导致科技论文从投稿到发表所需时间增加。由于核心期刊的稿源丰富,编审人员工作认真、编辑质量高,因此从收稿到发稿时间比较长,论文经编辑加工、修改后才发表有的需要两年之久。随着网络技术的普及,以电子邮件方式投稿、修回、出版发行已经成为期刊文献的编辑出版方式,目前国内外相当多的出版社已经采用这种全新的“即时出版”方式,当一本新杂志在电脑上编辑完毕,不再需要繁琐的出版周期,只要将其并入网络,几乎同时所有订户都能在终端屏幕上读到其内容。

四、文献检索的目的和意义

1. 继承和借鉴前人成果,避免重复研究 科学技术的发展具有连续性和继承性的特点。科学的每一进展都是在前人知识的基础上取得的,科学知识的每一次进步都是对某一方面的补充、修改和精练。对于一个科研工作者来说,文献信息检索是其科学研究不可缺少的一部分,在其着手开展一个新的研究课题之前,必须全面了解这个课题的背景资料,如前人在这方面已做了些什么工作,是怎么做的,还存在什么问题等等。只有准确地搜集了相关资料,摸清了该科学技术发展的水平、动向,吸取已有的科技成果,才能避免科研工作中重复浪费,少走弯路。例如美国 20 世纪 50 年代,为了做“继电器节点电路合成研究”,曾经联合几家实验室研究了 5 年,耗资 50 万美元终于完成,但当发表成果时才发现该项目早已有其他人完成。又如美国某轧钢厂一位化学家,花了 1 万美元完成了一系列实验,并解决了问题,而正当他为自己的成果得意时,却得知有一份德国人的报告与其实验完全相同,全部资料只需花 5 美元就行了。这些例子都说明,如果在课题研究之前,能够先检索有关文献,了解其当前发展水平,就一定能避免重复劳动,并扬长避短,开拓思路,在一个新的起点上作出努力。

2. 节省科研工作时间,提高科研效率 文献的庞大数量和迅速增长,以及文献的分

散、交叉重复加重了研究人员搜索信息的负担。许多科学家和信息学家一再表明,不会获取信息将对科学研究造成损失和失败;不会获取有用的信息,将让科研人员在五光十色的信息海洋中感到困惑和无奈。但是如果懂得文献检索的方法和技巧,尤其是以计算机检索为主体的现代化信息检索技巧,就可以比较容易检索到针对性较强的文献,达到事半功倍的效果。

3. 协助决策者做出正确决策 所谓决策,就是指为了实现某一特定目标,从客观实际出发,在占有一定信息和经验基础上借助了一定的方法,从各种可供选择的方案中选出最佳方案的活动。正确的决策虽然与多种因素有关,如决策体制、决策方法、决策者的能力和素质等,但是决定性的因素还是决策者对客观实际的了解,对未来行动和后果的正确判断等,而所有这些主要依赖于全面、即时、准确的信息保障。“知己知彼,百战不殆”,所以说,情报是决策的重要依据。

4. 终身学习、知识更新的必备工具 时代在发展,信息在瞬变,科学技术日新月异,知识更新周期缩短,尤其以计算机技术和卫星通讯技术为主的现代信息技术更是变幻莫测。21 世纪,知识的新陈代谢更是加快,学习、吸收和应用新知识、新信息、新技术的能力已构成竞争力的关键因素。与此同时,为了迎接知识不断更新所带来的挑战,人们需要从在校学习转化为终身学习,而终身学习、知识更新都依赖于信息的获取和利用。

5. 开发情报信息资源,提高经济效益 近十年来,世界在发生翻天覆地的变化,在目不暇接的经济万花筒背后,一种新的经济形式——知识经济正向我们席卷而来,它正将人类社会带向 21 世纪知识经济时代。知识经济是社会发展的必然。早在 20 世纪 70 年代,美国未来学者托夫勒在其著作《第三次浪潮》中就指出了人类将面临信息社会的到来,此后,经济学者们开始关注这一提法。回顾近半个世纪以来的世界经济发展状况,50 年代西德重新崛起,60 年代日本经济奇迹般增长,70 年代出现亚洲四小龙,进入 80 年代停滞徘徊多年的美国经济也开始复苏并得到发展,这些发展都是以知识为基础的。1990 年联合国研究专家正式提出了“知识经济”这一概念,知识成为发展经济的最重要的直接资源,而信息成为一个国家生存和发展的关键所在,成为国际竞争中的重要因素。控制了信息,尤其是高新尖端的科技信息,也就获得了经济发展的主动权和国际竞争的优势。在这样一个社会经济概念面前,世界各国、各民族、各行业都在积极应对。谁能有效地捕捉信息,分析研究信息并加以利用,谁便能在强手如林、变幻莫测的市场中抢占有利的位置。

第二节 文献检索基础知识

一、文献检索的概念、原理和意义

(一) 文献检索的相关概念

文献检索是将文献按照一定方式集中组织和存储起来,并按照文献用户需求查找出有

关文献或文献中包含的信息内容的过程。所以广义上的文献检索实质上是包括文献的存储和文献的检索两个过程。而我们通常所说的文献检索就是找出自己所需的文献,这只是对文献检索狭义上的理解。文献存储是标引人员对文献的有序化存储,即检索工具的编制(输入),由专职工作者来完成。将大量分散的文献收集起来,深入理解原文后对其进行加工,即将原文主要内容转换成特定检索语言(标引),最后编制成有序的、系统化的检索工具书或数据库。文献检索是检索工具的利用(输出),即读者利用。用户根据自己的检索需要,选定检索语言和相对应的检索途径,从检索工具书或数据库中查找到所需文献。文献的存储和文献的检索两过程是相辅相成、密不可分的。所以文献检索的全称又叫“文献存储和检索”。

文献检索的方式一般分为手工检索和计算机检索。手工检索方式是指利用印刷本检索工具进行人工查阅,这类检索一般不能实现全文检索服务。计算机检索是指利用计算机、激光或磁存储器以及现代化通信手段所实现的检索方式,如:光盘检索、联机检索和网络检索,这种检索方式为全文检索服务提供了有利条件。计算机检索以其数据更新快、检索速度快和智能化多途径检索等诸多优点,已成为现阶段文献检索的主要方式,但是计算机检索要完全替代手工检索还需要一段时间。计算机检索与手工检索的原理是一致的,在学习有关检索语言、索引技术和信息资源方面的知识时,手检比机检更容易理解。所以了解手工检索知识是很有必要的,是掌握计算机检索的必要基础。

(二) 文献检索的原理

文献检索包括文献的存储和检索这两个过程,文献检索的核心就是要使文献的存储与检索两过程所采用的特征标识达到一致。文献的存储过程就是检索工具的编制过程,这一过程通常由信息工作者去完成;而文献的检索过程则是检索工具的利用过程,可由文献需求者或情报人员去完成。所以任何检索工具都具有存储和检索两方面的职能。存储是检索的基础,而检索是实现存储的目的和手段,两者相辅相成、互为依存。

存储过程,就是形成文献特征标识,为检索提供有规可循的检索途径的过程。该过程首先是要为每一篇纳入检索工具(或系统)的文献标记上检索标识,即对文献进行全面分析,从自然语言转换成规范化的检索语言的过程,这也就是文献的标引过程。检索标识有两种:一种是文献的外表特征,也称为自然标识,包括标题、著者、来源、卷期、页码、年月、类型、号码、文种等项目;另一种是文献的内容特征,也称为人为标识,主要是文献论述的主题,亦即文献的中心内容,诸如主题词、分类号、类目名称、文摘等。标引内容特征时,首先需要对原始文献的内容进行主题分析、深入理解,找出若干能代表文献中心内容的概念,然后用情报检索语言的词语(或标识)来把这些概念标识出来;如对文献给予分类号的过程称为分类标引,给予主题词标识的过程称为主题标引。其次是要将有关文献特征标识著录下来,形成一条条有规可循的文献线索。最后将这些文献线索系统地按一定规则排列到检索工具(或系统)中去。

检索过程,就是利用特征标识,找出相关文献的过程。该过程首先是要对检索课题进行主题分析,找出若干能代表课题需要的概念,即确定主题检索概念;然后用存储过程中所使用的情报检索语言,将这些主题检索概念转换成检索特征标识;最后利用这些特征标

识到检索工具（或系统）中去查找到相关文献线索。

由此可知，检索过程是与存储过程相对应的逆过程，而连接两个过程的纽带就是情报检索语言。文献检索的基本原理就是检索者将主题检索概念与存储在检索工具（或系统）中的检索特征标识进行比较，最后达到一致，以此来有效地查得并获取文献。

二、检索语言

文献检索的原理是要使文献的存储与检索两过程所采用的特征标识达到一致，而使其一致的中间纽带就是检索语言。检索语言是检索者与检索工具之间进行交流的标准，也是情报存储人员与检索工具之间交流的标准。检索者、检索工具、情报存储人员三者通过检索语言来互相沟通。

（一）检索语言的概念

检索语言（retrieval language）是用于描述检索系统中文献的内容特征及外表特征和表达用户检索提问的一种专门语言。检索的匹配就是通过检索语言的匹配来实现的。检索语言是人与检索系统对话的基础。检索效率的高低，在很大程度上取决于所采用的检索语言的质量以及对它的使用是否正确。

（二）检索语言的类型及特点

检索语言由检索标目和语法组成。检索标目是指登录在检索系统或工具中的具有检索意义的文献特征标识，语法则是指利用这些标目查找文献的一整套规则。检索语言有很多种，按照不同的标准可有不同的划分，归纳起来主要有以下两种分法。

1. 按照检索语言中所使用语词的受控情况可分为规范语言（controlled language）和非规范语言（uncontrolled language）。

规范语言，又叫受控语言、人工语言（artificial language），是一种采用经过人工控制了的规范性的词语或符号作为检索标目，来专指或网罗相应的概念，这些规范化的标目能较好地对同义词、近义词、相关词、多义词及缩略词等概念进行规范。比如“肿瘤”这个概念的英文词汇中有多个同义词：tumor、tumour、carcinoma、neoplasms 等，人工语言则规定只能用其中一个词来表示所有“肿瘤”的概念，假如人为规定的词是“neoplasms”，那么无论原始文献中作者使用的是哪一个词，检索者只需用“neoplasms”就能将所有肿瘤的文献查找出来。因此，使用规范化的词能相对提高检索的效率，但对检索者和情报存储人员在选词上要求比较严格。在后面所述的单元词语言、标题词语言、叙词语言等都是规范性语言。

非规范语言，又叫非受控语言、自然语言（natural language），是采用未经人工控制的词语或符号来作为检索标目，通常所说的自由词、关键词就属此类。一般当某些特定概念无法用规范词准确表达，或新出现的词语（如纳米）还未来得及被规范化时，都需要使用

非规范词。这类词语有较大弹性和灵活性,检索者可以自拟词语进行检索。在机检中,非规范词的应用比较广泛。但这类语言对一词多义、多词一义的词语,检索就相对困难些。在后面所述的关键词语言就是一种非规范语言。

2. 按照检索语言所描述的文献信息特征的内容来分,其类型如图 1-1 所示。

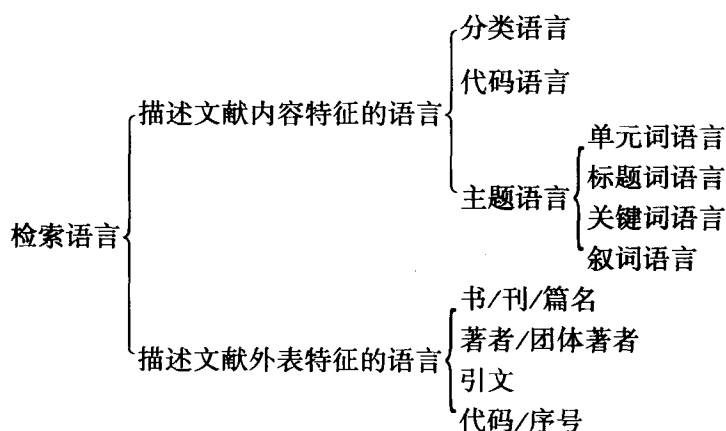


图 1-1 检索语言的类型

不同的检索语言构成不同的标识和索引系统,提供用户不同的检索点和检索途径。

描述文献外表特征的检索语言,是依据文献外表特征作为文献存储的标识和文献检索提问的出发点而设计的索引语言,如以文献上记载的书名、刊名、篇名等作为标目的书名目录(索引)、刊名目录(索引),以及“引用期刊一览表”、篇名索引等;以文献上署名的著者、译者、编者等姓名或机关团体名称作为标识的著者目录(索引)、专利权人索引等;还有以文献末尾所附的参考文献或引用文献的特征作为标识的,如引文索引;以及以文献特有的序号或代码作为标识的索引系统,如专利号索引、科技报告号索引等,这些都是采用描述文献外表特征的检索语言建立起来的。

描述文献内容特征的检索语言,按其构成原理分为分类语言、代码语言和主题语言。

(1) 分类语言 是一种直接体现知识分类的等级制概念标识系统,是按照学科范畴划分而构成的一种语言体系。它能集中体现学科的系统性,反映事物的从属、派生关系,并从上至下、从总体到局部层层划分、展开,是一种等级分明的语言。这种语言一般由类目(语言文字)或其相对应的类号(字母、数字或它们的组合)来表达各种概念,构成一个完整的分类类目录,如《中国图书馆分类法》、《国际专利分类表》等都是典型的分类语言。

(2) 代码语言 是对文献所论述事物的某一方面的特征,用代码加以描述和标引,如:化学物质的分子式、化学物质登记号、基因符号等。

(3) 主题语言 根据其表达概念的不同形式又分为关键词语言、标题词语言、单元词语言和叙词语言。

1) 关键词语言(keyword) 是由直接从文献的篇名、文摘或全文中抽取出来的词汇构成的。在检索工具中常以关键词索引出现。这些词汇都对表达文献主题具有一定的实质意义,几乎所有有意义的信息单元都可以作为关键词,而那些对文献主题表达无意义的词

汇,如冠词、介词、连词、副词、感叹词,以及某些形容词、某些动词(系动词、情态动词、助动词)等,均不能作为关键词,并被归入非关键词表或禁用词表。关键词语言基本上是未经规范化处理的语言,属于自然语言的一种,但实际运用中仍对少量词作了规范化处理,如复合词的拆散和组合,词和符号的改写等。关键词语言具有易于掌握,方便检索,灵活性高,在计算机检索中应用广泛等诸多优点,但它未经规范,用词不统一,因而漏检率高,检全率较低。

2) 标题词语言(heading) 是最早使用的一种主题词语言,它使用的词汇也是从文献内容中抽出,但与关键词语言不同的是,它们都是能代表文献内容特征的规范化词语,属于规范性语言的一种。

标题词分为主标题词和副标题词。标题词表中出现的这两级标题词以固定的方式组合在词表中,检索按既定组配进行。因而使用标题词语言检索之前,词与词之间的组配已经固定,这种形式又称为先组式语言。标题词语言属于规范性语言,因而具有较好的通用性、直接性和专指性,但其又是先组式语言,因此灵活性较差。

3) 单元词语言(uniterm) 是一种最基本的、不能再分的单元词语,亦称元词。它也是从文献内容中抽出,并经过规范化处理的。但它仅能作为元词(或单元词),代表一个独立的概念,比如“肝功能”不是单元词,只有“肝”和“功能”才是单元词。在英语中元词经常是一个单词。元词具有灵活的组配功能,在检索时可以将某些元词组配起来代表某一特定概念,因此这种语言属于后组式语言,如《世界专利索引》工具中的《WPI-规范化主题词表》就属于单元词语言。

4) 叙词语言(thesaurus) 叙词是在标题词、单元词和关键词基础上发展起来的,它和关键词语言是目前使用较多的两种检索语言。叙词语言是规范性后组式语言,与单元词语言不同的是,它既包括了单个的词,也包括了词组,它可以用复合词表达主题概念,如“肝功能”就可作为一个叙词;而且单元词是字面组配,叙词则是概念组配,如“香蕉苹果”这个概念,在尚未成为叙词和单元词时,单元词语言组配为“香蕉+苹果”,而叙词语言组配为“带有香蕉味的水果+苹果”。美国《医学索引》中的《医学主题词表》就是典型的叙词语言的一种。

综上所述,叙词语言集规范性语言与后组式语言之优势,既在用词上达到统一,又有较好的灵活性,与其他类型的检索语言相比,它是一种多功能的较完善的检索语言。科技人员都熟悉自己的专业词汇,从叙词表中找到切题的叙词后,组配得当,就能大大提高检索效率。

三、检索工具

(一) 文献信息检索工具的概念

文献信息检索工具,是指用以存储、报道和检索文献的工具。文献信息检索工具都要收摘报道一定范围的一次文献;对所收摘文献的外表特征和内容特征进行描述,并且对描