

21 世纪高等学校精品规划教材

医学信息检索

第二版

刘薇薇 王虹菲 方立 主编



天津大学出版社

21 世纪高等学校精品规划教材

医学信息检索

(第二版)

主 编	刘薇薇	王虹菲	方 立
副主编	赵云艳	曾家琳	
参 编	邓岩彬	冯 研	刘 强
	吕晓妍	马子龙	王 馨
	王 宇	夏雪雁	许竹萍
	杨 峰	杨丽娟	叶 翎
	原 增	张 蕾	赵志强

内 容 提 要

本书是针对信息检索数据库和网络信息资源的变化特点,围绕“医学信息检索”课程的目标与内容编写而成的,主要包括信息检索的基础知识、中外文医学信息检索数据库的检索与利用、相关信息资源的检索与利用,以及图书馆利用的基本知识和联合目录的检索等内容。本书详细、系统地介绍了检索原理与技巧,并辅以检索实例,可操作性强。本书既可以作为医学专业本科生、研究生和七年制学生的“医学信息检索”课程教材,也可以作为医学专业人员学习医学信息检索知识、提高检索技能的参考工具书。

图书在版编目(CIP)数据

医学信息检索 / 刘薇薇,王虹菲,方立主编. —2版. —天津: 天津大学出版社, 2009 (2011.8重印)

21世纪高等学校精品规划教材

ISBN 978-7-5618-3220-2

I. ①医… II. ①刘…②王…③方… III. ①情报检索

IV. ①G252.7

中国版本图书馆CIP数据核字(2009)第164535号

出版发行	天津大学出版社
出 版 人	杨欢
地 址	天津市卫津路92号天津大学内(邮编: 300072)
电 话	发行部: 022-27403647 邮购部: 022-27402742
网 址	www.tjup.com
印 刷	北京紫瑞利印刷有限公司
经 销	全国各地新华书店
开 本	185mm×260mm
印 张	11.5
字 数	288千
版 次	2009年9月第1版 2011年8月第2版
印 次	2011年8月第2次
定 价	20.00元

凡购本书,如有缺页、倒页、脱页等质量问题,请向我社发行部联系调换

版权所有 侵权必究

前言

信息素质是现代人才必备的基本素质之一，是人们终身学习的基础。大学生信息素质培养已成为新世纪人才培养的重要内容。信息检索与利用课程是一门实践性强、内容更新比较快的方法性课程，旨在培养学生的信息意识和检索技能，提高信息利用能力，是培养学生信息素质的重要途径。

自 1984 年教育部颁发关于在高等学校开设“文献信息检索与利用”课程的文件以来，全国各类高校相继将该课程纳入学校课程体系。医学院校开设的“医学信息检索与利用”课程，顺应信息时代发展的需求，在培养医学专业学生信息素质方面发挥了积极的作用，因而表现出旺盛的生命力。随着信息技术的发展和网络应用的普及，“医学信息检索与利用”的教学内容不断更新与丰富，相应的教材也可以说是百花齐放，各具特色。我们结合多年来教学理论的研究与信息检索的实践经验，紧紧把握医学信息检索技术发展的脉搏，以信息检索基础知识与基本方法为基础，突出检索要点与技巧，以检索实例为核心，分析各检索系统的特点，帮助学生理解和掌握各检索系统的使用方法，具有很强的实用性。在第一版的基础上本书又增加了文献管理一章，以帮助学生进行有效的文献管理。

本书可以作为医学本科生、研究生、七年制学生的“医学信息检索”课程教材，也可以作为医学专业人员学习医学信息检索知识、提高检索技能的参考工具书。由于编写时间紧迫和知识所限，书中难免存在错误与疏漏之处，特别是网络信息变化更新很快，书中信息难免存在过时现象，敬请广大读者批评指正。

编者

目 录

第一章 概论	1
第一节 文献	1
第二节 检索的基本原理与检索语言	5
第三节 医学文献检索常用的检索语言	9
第四节 计算机检索技术	13
第五节 检索工具与检索步骤	15
第二章 中外文医学文献检索工具	20
第一节 中国生物医学文献服务系统 (SinoMed)	20
第二节 MEDLINE 数据库	28
第三节 PubMed 检索系统	42
第四节 BIOSIS Previews	52
第五节 International Pharmaceutical Abstracts	57
第六节 国家科技图书文献中心	58
第三章 中文医学文献全文数据库	61
第一节 中国知网 (CNKI)	61
第二节 维普期刊资源整合服务平台	69
第三节 万方数据资源系统	72
第四章 英文医学文献全文数据库	77
第一节 ScienceDirect	77
第二节 OvidSP 期刊全文数据库	81
第三节 WILEY ONLINE LIBRARY	85
第四节 SpringerLink	88
第五节 其他外文全文数据库	90
第六节 网络免费期刊全文资源	93
第五章 其他类型数据库	98
第一节 超星数字图书馆	98

第二节	读秀中文学术搜索	102
第三节	apabi 数字资源平台	104
第四节	书生之家	108
第五节	外文电子图书	109
第六节	中国疾病知识总库	114
第七节	医学考试培训系统	116
第八节	多媒体视听资料的获取与利用	118
第六章	网络信息资源	120
第一节	搜索引擎	120
第二节	基础医学网络资源	127
第三节	临床医学网络资源	130
第四节	循证医学信息资源	135
第五节	生物信息学数据库	140
第七章	图书馆利用	146
第一节	图书馆功能介绍	146
第二节	如何利用图书馆	148
第八章	文献资源管理	154
第一节	EndNote	154
第二节	其他文献管理软件介绍	159
附录 1	《中国图书馆分类法》(第 5 版) 简表	161
附表 2	MeSH 医学副主题词使用范围一览表	166
附表 3	副主题词类别一览表	175
附录 4	常用网址	176
参考文献		178



第一章 概论

无论是日常生活、学习工作，还是科学研究，人们都离不开文献信息资源的支持。对医学专业的学生来说，既强调知识和经验的积累，也强调知识更新和科学证据。在临床实践中，医生作决策时，除运用自己的临床技能和治病经验外，也常常依赖相应的医学研究文献以佐证、完善或纠正自己的判断。随着现代科学技术的迅速发展，生物医学文献呈现出增长迅速、数量庞大、类型众多、获取方式复杂、评判难度加大的特点。在浩大的信息资源中决定所需信息的范围，有效地获取所需信息，严格评价已获知的信息及其相关资源，并进行系统的分析整理，这是进行临床医疗或其他医学相关工作的重要能力之一。

医学信息检索与利用是一门科学方法课，专门研究各种生物医学信息资源，包括图书、期刊和其他出版物、专业数据库以及互联网上相关资源的获取与利用，注重培养医学及相关专业学生的信息素养和检索技能。提高信息获取、选择、分析与利用的能力，能够把所获信息有机融合到个人的知识库中，有效运用信息达到特定目的，对于避免未来工作和科学研究中走弯路、解决实际问题、提高创新开拓能力以及终身自我学习、更新专业领域知识都具有十分重要的意义。

第一节 文献

一、文献的概念

文献是记录有知识的一切载体。即是说，文献是将知识和信息的内容用语言、文字、符号、图像、音频、视频等方式记录在纸张、胶卷、磁带、磁盘、光盘等各种介质载体上的结合体。文献包含知识内容、载体形态、记录方式三个要素，具有存储知识、传递信息和交流信息的功能。

信息是人们对客观世界中各种事物的变化状态和规律特征的反映，是客观事物经过感知或认识后的再现，具有主观和客观的两重性。客观世界中不同的事物是千差万别的，呈现出不同的状态和特征，从而在人们认识的过程中形成不同的信息。信息分为两大类：自然信息（包括生物信息）与社会信息（包括机械信息），能够通过感觉、文字、图像、声音、符号或数据等为人类所获知，被传递、被共享。

知识是人类对客观事物的认识和实践经验的总结。它对信息进行提炼和深化，形成较为系统、完整的科学体系，是人们科学地认识世界、改造世界的力量来源。情报来源于知识，

必须在特定的时间内经过及时传递,并能为特定用户所接受和利用,既可以是口头传述,也可以是以文献形式记录,具有知识性、传递性和效用性。知识被记录形成文献,被传递成为情报。它们之间的关系如图 1-1-1 所示。

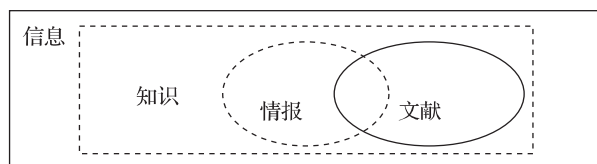


图 1-1-1 文献、信息、知识与情报的关系

信息与知识是构成文献的主体。没有信息或知识内容的任何形式的载体,不能称为文献。任何形式、任何类型的文献都必须首先以具有一定的信息和知识内容为前提,用某种信息符号记录于某种载体上,即信息和知识内容与所附载体是不可分割的。文献记录了人类同自然抗争的历史和社会发展的进程。医学文献则是人类长期同疾病作斗争的经验总结和对生物医学研究的知识积累。人类社会利用文献或通过文献进行交流,其实质就是利用文献中的信息、知识和情报。知识的积累和信息的传播,是人类知识发展和文明进步的基础。而网络的存在与普及、媒介的多样化以及数据挖掘技术的发展,使得文献信息传播得更为迅速和广泛,影响和作用越来越大。

1. 文献的作用

1) 文献是构成人类知识宝库的一部分,是记录、存储、积累和传递科技知识的重要工具,是传播、交流情报和更新知识的主要信息载体。

2) 文献是开展科学技术研究的依据和评审、确认先进成果的基础,可作为科研开题立项或确认其作者对科研发现、发明创造是否具有优先权的基本依据。

3) 发表文献的数量是衡量科技人员创造性劳动效率的指标之一,也是衡量一个科学机构或国家地区的学术水平的重要标志。

4) 个人通过对某个专题文献进行全面的搜集分析,能够使自己的知识库得到拓展和深化。

2. 文献的发展特点

1) 数量增长迅速,出版类型复杂化。

2) 由传统的纸本印刷型迅速地向电子型、网络化、数字化方向发展。

3) 出版发表的内容重复又分散,出版周期缩短。

4) 纸本出版物发表滞后期长。

5) 时效性增强。

6) 文种多,但英文文献占较大比重,非英文文献则越来越多地带有英文摘要。

7) 知识分类清晰,专业交叉渗透,主题词数量增长快。

8) 借助网络化和数字化,信息传播与共享、数据内容检索更加便利。

9) 全文服务功能更加全面,检索平台越来越注重个性化服务。

10) 开放存取资源发展较快, 实行同行评议控制文献质量。

二、文献的类型

1. 按文献载体类型划分

为了有效地存储、传播知识, 人类先后发明了各种各样的物质材料来记录信息。随着信息记录方式与存取技术的发展, 文献的范围进一步扩大, 文献的生产和传递更加迅速, 知识、信息的存储和利用更加便捷, 文献载体形式呈现多样化。

(1) 纸本型 (书写型、印刷型)

纸本型文献是以手写、雕刻、打印、印刷等为记录手段, 将信息记载在纸张上形成的文献, 如手稿、笔记、图书、期刊等。

(2) 缩微型

缩微型文献是利用光学技术以缩微照相为记录手段, 将信息记载在感光材料上形成的文献, 如缩微胶卷、缩微平片等。

(3) 声像型 (视听型)

声像型文献是采用录音、录像、摄影、摄像等手段, 将声音、图像等多媒体信息记录在光学或磁性材料上形成的文献, 如音像带、唱片、幻灯片等。

(4) 电子型 (机读型)

电子型文献是指以数字代码方式将图文、声像等信息存储到磁、光、电介质上, 通过计算机或类似设备阅读使用的文献, 如电子图书、电子期刊、数据库、网络新闻等。

目前, 为了避免因版权所有者或物理存储地址等属性的变更而改变网址, 许多电子文献或数据库中的电子全文标上了 DOI (Digital Object Identifier, 数字化对象识别符), 具有唯一性、永久性、通用性, 能够保证在网络环境下对数字化对象的准确提取。目前国际上有一种合作式参考文献链接服务——CrossRef, 就是利用了 DOI 来实现不同出版商出版的在线学术资源之间高效而可靠的交叉链接, 可通过点击期刊论文的一条参考文献引用条目来访问被引用的文章。

2. 按文献出版形式划分

(1) 图书

图书是指对某一领域的知识进行系统阐述或对已有研究成果、技术、经验等进行归纳、概括的出版物。图书内容系统全面, 知识可靠成熟。按照内容、著述目的和使用特点, 一般可分为普通图书和工具书。普通图书又称阅读用书, 包括教科书、专著、文学艺术书、专辑、汇编、论丛、多卷书和丛书等。工具书主要分参考工具书和检索工具书, 包括指南、百科全书、年鉴、手册、辞典、统计数据汇编、地图、图谱、人名录及索引、目录、文摘等。

识别图书的主要依据有: 书名、著者、出版地、出版社、出版时间、版次、总页数、国际标准书号 (ISBN) 等。

(2) 期刊

期刊是指有固定名称、版式和连续编号, 每一期可刊载由不同著者撰写的多篇文章或论文, 装订成册, 定期或不定期地长期出版的连续性出版物。特点是内容新颖、信息量大、出

版周期短、传递信息快、传播面广、时效性强,能及时反映国内外各学科领域的发展动态。

期刊按内容和性质划分,主要有由学术团体编辑出版,报道生产、科研方面的学术论文及研究成果等,如学术性期刊(学报、通报或公报、汇刊、评论、进展等),快报性期刊(研究通讯、新闻短讯等),获取文献线索的检索性期刊(文摘、索引),保存历史记录的资料性期刊(资料、事件)等。

所谓专业核心期刊是指刊载该专业论文数量较大,学术水平较高,能反映本学科最新研究成果及本学科前沿研究状况与发展趋势的,备受该学科专业读者重视,影响因子系数高的学术期刊。每个学科专业都有自己的核心期刊。核心期刊与非核心期刊的划分是相对的、动态变化的。国内认定核心期刊的工具较多,最常用的工具书有北京大学编制的《中文核心期刊要目总览》(每四年修订1次,目前最新版是2008年版,其中有生物医学核心期刊200余种)、中国科学技术信息研究所每年出版的《中国科技期刊引证报告(核心版)》的统计源期刊目录等。国际上的科技核心期刊则通用美国科学情报研究所(ISI)的 *Journal Citation Reports* (JCR, 期刊引用报告)。

识别期刊的主要依据有:名称,出版的年、卷、期,国际标准刊号(ISSN)等。期刊在多年连续出版的过程中,常常有停刊或复刊、合并或分出等情况。国外出版社把丛书(Series)视作连续出版物,常使用ISSN号。

(3) 报纸

报纸也是一种连续出版物,但一般以新闻与时事评论为主要内容,每期有多个版块(主题),很少装订成册,如日报、周报等。

(4) 特种文献

特种文献指图书、报刊之外,出版发行和获取途径都比较特殊的科技文献,包括科技报告、专利文献、标准、会议资料、学位论文、政府出版物、产品样本和说明书、住院病历等。

3. 按文献级次划分

(1) 一次文献

一次文献是人们直接以自己的劳动、科研、社会活动等实践经验为依据生产出来的,具有一定发明创造或一些新见解的原始文献,包括期刊论文、专著、专利文献、科技报告、会议录、学位论文、技术标准,等等,其所记载的知识、信息比较新颖、具体、详尽,在整个文献中是数量最大、分布最分散、种类最多、使用最广、影响最大的文献,具有创造性、原始性、多样性等特点,是最主要的文献信息源和检索对象。

(2) 二次文献

二次文献是指在特定的学科范围、出版时间或出版类型内,将大量分散、零乱、无序的一次文献线索按其特征(著者、篇名、分类号、出处、关键词、主题词等)进行整理、提炼,并按照一定的逻辑顺序和科学体系加以编排存储,以便于检索利用的文献。其主要类型有书目、题录(包含有题名、作者、作者机构、文献来源及其他简要信息)、目录、文摘、索引、书目性数据库、网络学科资源导航、搜索引擎等。二次文献收录的不是一次文献本身,而是一次文献线索,具有明显的汇集性、系统性和工具性。

(3) 三次文献

三次文献通常是指围绕某个专题利用二次文献进行检索搜集,从中选用大量相关文献,

经过综合、分析、研究等一系列深度加工而编写出来的文献,包括综述、述评、进展、动态、参考工具书(如:辞典、百科全书)、文献指南等,具有综合性、针对性、科学性等特点,指引性强,内容浓缩,信息量大。在科学研究工作中,利用三次文献有助于在短时间内了解具体的研究历史、发展动态等,更全面、准确地掌握待查课题和项目的技术背景与研究进展。

还有一种特殊形式的没有经过正式发表的情报信息源,有时被称作“零次文献”,主要包括两类:①形成一次文献以前的知识信息,即未经记录,未形成文字材料,人们口头交谈的情报信息;②未经正式发表的原始材料,如书信、手稿、设计草图、实验记录、发言稿、会议记录、笔记等一些内部使用且不能通过公开的途径所获得的资料。其特点是内容新颖,不太成熟,一般是通过口头交谈、参观、与会等途径获取,目前越来越多地借助网络进行非正式的传播。例如,病例报告、实验方法,近年常常可以在网上的专业论坛中见到,有的记录内容相当完整。这些信息资源不仅在内容上有一定的价值,而且新颖程度高,但很少像正式文献那样公开引据作为参考文献。

从文献加工生产过程可以看出,从零次文献源到一次文献、二次文献、三次文献,是一个由分散到集中、由无序到有序、由博到精地对知识信息进行加工的过程。零次信息是最原始的信息资源,虽未正式公开交流,但它是生成一次文献的主要素材;一次文献是最基本的信息源,是检索和利用的主要对象;二次文献是对一次文献的搜集和有序化,是检索的工具;三次文献是把一次文献,按照专题或知识的门类进行综合分析加工,高度浓缩成新的信息源,既是检索和利用的对象,又可作为检索工具去检索、获知和获取一次文献内容。随着全文索引技术及海量存储技术的发展,全文数据库已集一次文献的存储、二次文献的检索、三次文献的专题查阅功能于一体,极大地方便了用户的信息利用。

第二节 检索的基本原理与检索语言

文献检索是使用科学的方法,从不同信息集合中迅速、准确、完整地查找和获得文献资料。这里所说的信息集合,是指有组织的信息资源系统,可以是检索性工具书、资源门户,也可以是各种数据库,如图书馆馆藏书目检索系统、中国知网、PubMed,等等。

一、检索基本原理

检索基本原理就是将用户的提问特征与检索系统的文献标识特征进行对比,并将其中匹配一致或比较一致的数据输出给用户。

为了实现标识特征的对比检索,需要分析出所收录(存储)的每个文献的外部特征和内容特征,用约定的检索语言进行标引,按照规定的方式组织到编制好的检索系统里;检索时需要根据检索者的特定需求,确定提问特征,从检索系统中查找出有关文献。文献检索过程如图1-2-1所示。

从图1-2-1可以看出,文献检索包含了两个过程:

1) 存储过程: 通过特征分析, 使用检索语言标引记录和组织索引, 把大量分散的无序的文献集中存储起来, 成为具有多种索引功能的检索系统;

2) 检索过程: 利用已组织好的检索系统, 按照检索要求使用检索语言将提问标识与检索系统中存储文献的特征标识进行人工比较, 或是通过计算机进行字符串匹配和逻辑运算, 输出命中文献, 从而将所需信息查找出来。因此, 信息检索的全名为信息存储与检索, 广义的概念是指存储与检索两个过程, 狭义的信息检索仅是指检索过程。

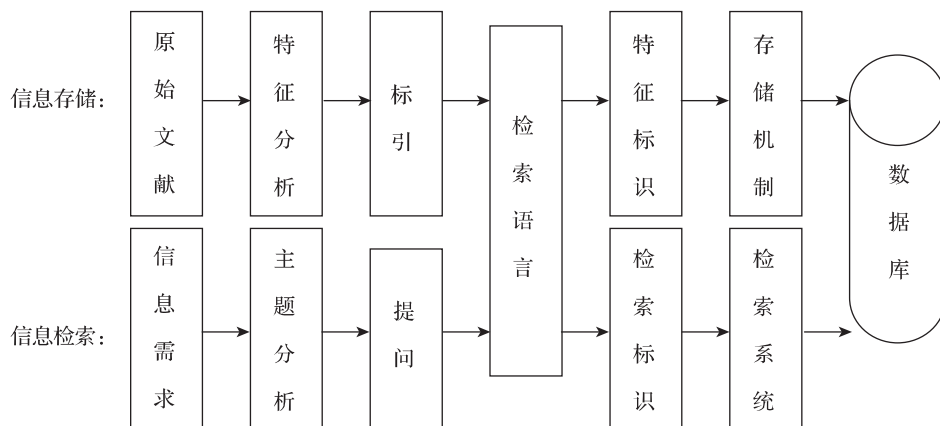


图 1-2-1 文献检索过程图

依据检索目的的不同, 文献检索可以分为以下几种类型:

- 1) 数据 (数值) 检索: 以数据、公式或数据型图表为查找对象, 如人体生理常数、血液检验正常参考值等;
- 2) 文献检索: 以文献为对象, 查找有关某个课题的原始文献或三次文献;
- 3) 图像信息检索: 以特定图像为查找对象, 如医学影像、电子地图等;
- 4) 事实检索: 检索某个特定事物或事件的始末、实验方法、试剂品牌等, 如检索 “川崎病” 发生的由来与治疗方案比较。

二、检索语言

从上面的文献检索过程可以看出, 检索语言起着极其重要的作用, 它是沟通文献存储与文献检索两个过程的桥梁。为有效地获得和利用文献, 人们必须通过检索语言使存储过程中的标引用词与检索过程中的提问用词达到一致。当用户的检索标识与文献的特征标识完全匹配或部分匹配时, 结果即为命中文献。

文献检索语言是用来描述文献特征和表达检索提问, 并为检索系统提供一种统一的符号化或语词化的专用人工语言。检索语言因其使用的场合不同, 也有不同的叫法: 用于文献的标引和存储的, 叫标引语言; 用于组织文献的索引的, 叫索引语言; 在检索文献过程中则称为检索语言。在检索系统中, 每种检索语言都对应一种检索途径 (检索点)。正确地运用检索语言有助于提高文献检索效率。

1. 以文献外部特征为检索标识的检索语言

文献外表特征有题名、著者、文献来源、期刊卷期、页码、出版年月、文种、文献类型、著者单位、参考文献等,可作为文献标识与检索依据,直接明了。相应地,则有题名、著者、文献出处(来源)、作者单位等检索途径。

(1) 题名语言

按文献的完整题目(题名、篇名)或题目中的关键词和词组进行检索的索引语言。

(2) 著者语言

按著者、译者、编者等姓名进行检索的索引语言。通用原则是:姓在前,名在后,姓用全称,名用缩写。例如,Michala E. Pedersen,通常使用“Pedersen ME”或用“Pedersen M”检索。检索英文文献时,中国著者用汉语拼音,日本著者用日本黑本式《日文字母-拉丁字母音译对照表》。具体规定见各数据库的检索说明。

(3) 号码语言

按文献的某个特定号码,如专利号、标准号、报告号、ISBN号、ISSN号等的顺序排检的索引语言。

(4) 引文语言

利用文献之间引用与被引用的关系,作为文献内容主题标识,并以此标引和检索文献的索引语言。标引词可以取自参考文献(即引文)中的题名、著者、来源等。引文检索既能提供同类研究的历史轨迹,也可通过被引用的情况反映该论文及其相关工作的学术水平,已被列为科技查新的一项重要内容。

2. 以文献内容特征为检索标识的检索语言

这类检索语言普遍应用了概念逻辑的原理来表达各种主题概念及其相互关系(等同关系、属分关系、相关关系),并且利用了“概念的划分与概括”和“概念的分析与综合”这两种逻辑方法来建立各自的结构体系。前者以等级体系分类法为代表,后者以叙词语言(主题词表)为代表。

文献内容特征可以是学科分类、主题内容、物质结构符号等,以其为检索标识的检索语言包括分类语言、主题语言、代码语言等,相应地有分类名称或分类号、主题词、关键词、代码等检索途径。

(1) 分类语言

广泛用于图书、资料的分类和检索,具体体现形式就是图书分类法。比较有影响的有《美国国会图书馆分类法》、《杜威分类法》和《中国图书馆分类法》等。它们用固定的分类号和相应的分类名称来表达各种概念,以科学知识分类为基础,结合文献信息的内容及其特征,运用概念划分和归属的方法,采取从总到分、从一般到个别、从抽象到具体、从低级到高级、从简单到复杂的原则进行层层划分,是等级体系结构。同位概念之间体现的是平行关系,而上下位类之间则是隶属关系。分类语言便于按学科门类进行族性检索,但不适用于多学科性的、按专题概念进行的特性检索。《国际十进分类法》在等级列举制的基础上采取了多种符号进行组配,能够揭示复杂主题的文献,因而发展成一部概念等级列举与分面组配相结合的混合式分类法,成为国际通用的情报检索语言之一。专利文献则使用专门的专利分

类法,国际上通行使用“国际专利分类法”(IPC)。

(2) 主题语言

主题语言将从文献中优选出来的,能够准确表达文献主题内容特征的词汇作为概念检索标识,按字顺排列,并用“参照系统”等方法辅助显示概念之间的相互关系,具有直观性强、专指性高、灵活性好、概括全面的优点。按照选词原则、词规范化处理的不同,主题语言包括标题词语言、单元词语言、叙词语言、关键词语言等。其中,关键词语言和叙词语言是目前使用较多的主题检索语言。

1) 关键词语言:关键词又称为自由词,是由直接从文献的篇名、关键词、摘要或全文中抽取出来的且能揭示文献题名或主要内容的关键性词汇构成的,未经规范性处理,检索信息比较直观,不受主题词表与学科体系的局限。几乎每一个搜索引擎或全文数据库都具有关键词检索功能。

2) 叙词语言:叙词是指一些以概念为基础的,具有组配功能,经过规范化的词或词组,并配有一个完整系统的主题词表。通过词表将各种同义词、近义词、多义词、同形异义词、先组式词组(在主题词表内事先将两个或多个主题词固定为一个词组)等严格规范化,并且揭示各词之间的语义参照关系,如同义关系、上下关系、相关关系。叙词主题词表经常增订,在检索词的统一和概念的组配有较好的灵活性,对一些新兴学科、边缘学科的查找尤其方便,能突破知识分类的框架,把分散在各学科里的有关课题的文献集中于同一主题下,有利于综合性检索。

(3) 代码语言

代码语言是对文献所论述事物的某一方面的特征,用某种代码系统加以描述和标引,如化学物质的分子式、化学物质登记号、基因代码等。这种检索语言尤其适用于生物化学、医学检验等专业的文献检索。

网络环境下还出现了另一类检索语言:网络分类法,如分类指南、学科导航。许多搜索引擎、数据库、学科门户、网站根据自身收录对象和用户检索特点自行设置学科大类类目,直接使用学科名称,不用类号,没有严格的等级划分,能基本满足用户的主题检索要求。

三、禁用词

禁用词(Stop Words),也叫停用词,即计算机检索中的虚词、非检索用词。有些词因为单独使用时无实际意义,如英文 and、or、not、in、at、of、about、up、out、is、are、the 等以及中文的连词、副词、介词、语气助词等虚词,或者出现频率过高,如中文“我们”,或者某些词及符号已用于系统中代表逻辑算符,就被定义为禁用词,不能作为检索词。若使用带有禁用词的语句进行全文检索,检索系统或搜索引擎将自动以空格方式忽略这些禁用词,继续执行其余检索词的检索。规定禁用词有利于节省索引空间和提高搜索效率。每个数据库都有自己的禁用词规定。

第三节 医学文献检索常用的检索语言

检索国内生物医学文献数据库最常用的检索语言工具有《中国图书馆分类法》、美国《医学主题词表》(MeSH)、《中国中医药学主题词表》，国外生物医学文献数据库则以MeSH为主。

一、《中国图书馆分类法》

《中国图书馆分类法》简称《中图法》，各图书馆多用第4版（第5版已于2010年8月出版）。它是以科学分类和知识分类为基础，并结合文献内容特点及形式特征进行逻辑划分和系统排列的类目表，采取从总到分、从一般到具体的逻辑系统，把知识门类分为马列主义、毛泽东思想、哲学、社会科学、自然科学、综合性图书五大部类，并在此基础上建成由22个基本大类（一级大类）组成的分类体系，用汉语拼音字母顺序来表示，约4万个类目，还有地区、时代、体裁形式等各种专用复分表。医药、卫生大类用字母“R”表示，有17个二级类目（详见附录1），各级类目下还可细分。类目级别可以根据类目号的位数（包括字母、数字，不包括中间的圆点）识别出来。例如：

一级类目	R	医药、卫生
二级类目	R5	内科学
三级类目	R54	心脏、血管（循环系）疾病
四级类目	R541	心脏疾病
五级类目	R541.1	先天性心脏血管病
		

图书馆书目检索系统（OPAC）和一些中文数据库都依据《中图法》提供详细的分类导航，可以通过点击逐级展开更专指、更具体的类目。图书馆的每一本藏书和多数国内学术期刊上发表的论文都单独标引有《中图法》分类号。大型文献类数据库都会对记录进行分类标引。

二、《医学主题词表》(MeSH)

医学主题词表（Medical Subject Headings, MeSH）由美国国立医学图书馆（NLM）编辑出版，是一个用以分析生物医学方面的期刊文献、图书、视听资料、电子资源等主题内容的叙词表，它也是NLM出版的MEDLINE/PubMED数据库主题检索的索引词典，原出版有纸本的词表，每年更新一次，自2009年起，纸本词表不再出版，而被电子版MeSH Browsers（主题词表查询器，在线主题词表，<http://www.nlm.nih.gov/mesh/MBrowser.html>）取代，每周更新，检索更为方便。MeSH有输入词语或词根前缀（Enter term or the beginning of any root fragments）和树状结构导航查询（Navigate from tree top）两种方式（查询界面如图1-3-1

所示), 查询结果也更为直观, 分两个部分: 上半部分是表格, 给出每一个主题词的详细注释说明 (主题词、树状结构号、词义注释、范围注释、入口词、参照主题词、允许组配使用的副主题词、历史注释等); 下半部分是相应的树状结构表, 显示出该主题词在树状结构中的位置、树状结构号及其上下位关系, 可以通过点击主题词树状结构号后面带的“+”号展开其下位主题词。

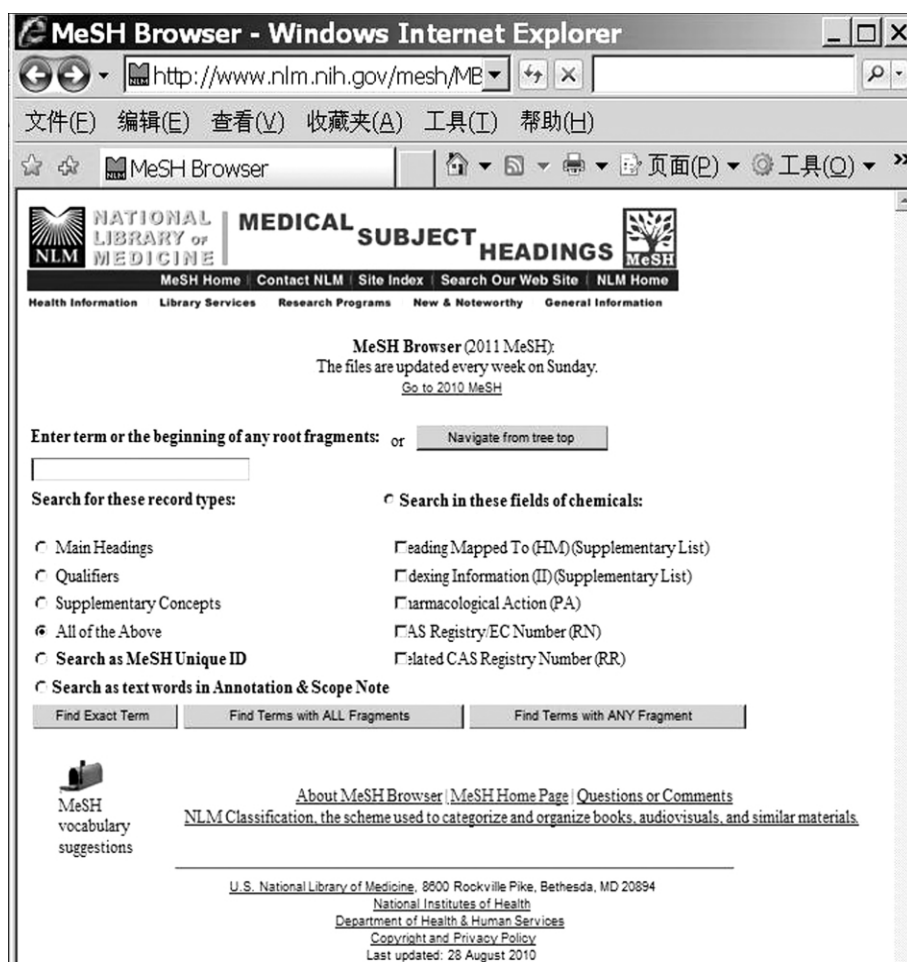


图 1-3-1 MeSH 查询器界面

MeSH 的词汇体系包括以下几种类型。

1. 叙词

叙词有以下三种基本类型。

(1) Main Headings (主题词, 即 Descriptors)

Main Headings 用以描述文献中心主题或内容特征, 每年修订, 2011 年比上一年增加近 600 个, 已有 26 140 多个。主题词以美式拼写为准, 一般使用正常的词序, 但为了使概念相近的词汇汇集在一起, 有时也采用倒置词序, 以便这些同族词能集中排列。例如:

Shock (休克)

Shock, Cardiogenic (休克, 心源性)

Shock, Hemorrhagic (休克, 出血性)

Shock, Surgical (休克, 手术性)

Shock, Traumatic (休克, 创伤性)

(2) Qualifiers (限制词, 即与 Descriptors 配合使用的副主题词)

每年更新, 但变化基本不大, 现有 83 个, 用于论述主题某一方面的内容, 对某一主题词的概念进行限定或细分。一篇生物医学文献论及的事物与研究的对象常常只涉及该事物的某个方面或某些方面, 如器官的解剖学、组织学、生理、病理等, 疾病的病因、病理生理、诊断、治疗等, 药物的治疗应用、副作用、药理学等, 通过主题词/副主题词的组配而成为一个更专指的概念。同时, 可以揭示主题词之间的逻辑关系, 避免词与词之间的虚假组配和语法歧义。每一个副主题词有特定的含义和使用范围, 必须与特定大类 (用树形编号表示) 的主题词相组配, 见附录 2。和主题词一样, 副主题词之间也有上下隶属关系, 见附录 3。由于副主题词范畴设置得很细, 检索时, 一个主题词可分别选用一个或多个副主题词进行组配, 或使用全部可用的副主题词, 以防漏检。

MeSH 查询器也能查询副主题词, 包括范围注释、词义注释、缩写形式、同义词等信息。当输入的检索词为主题词和副主题词都有记录时, 系统将会提示选择主题词还是副主题词检索浏览。

(3) Supplementary Concept Records (增补概念记录, SCRs)

以前称为“增补化学记录”, 每周更新, 现已有近 20 万条记录, 包括 MEDLINE 的化学药品、药物原始记录和小部分实验室指南, 具有独立的词库, 在 PubMed 数据库的物质名称字段 [NM] 检索使用。从 2011 年起, 库内还增加了罕见疾病增补概念记录。与主题词不同, 它们没有树状结构号, 但它们可与一个或多个主题词相连, 能够在 MeSH 查询器中查到, 还能看到有关化学方面的如药理作用、化学物质登记号 (CAS) 等的信息。

2. 参照

(1) Entry Term

入口词, 也称款目词, 不用作正式主题词, 而是指向 “MeSH Heading”, 用以协助查找准确的主题词, 2011 年有 17.7 万个。一般在以下几种情况下使用: ① 同义词和近义词, 如 Neoplasms 和 Cancer; ② 全称与简称, 如 Acquired Immunodeficiency Syndrome 和 AIDS; ③ 物质学名与俗名, 如 Ascorbic Acid 和 Vitamin C; ④ 疾病学名与俗名, 如 Glycogen Storage Disease Type IV 和 Andersen Disease; ⑤ 异形词, 如 Kidney Calculi 和 Renal Calculi。此参照常被称为 See cross-references, 相当于“代参照”(see 与 X)。

(2) See Also

提示与该主题词有关联的其他主题词。它们不属同一分类细目, 但彼此又有某种相关关系, 可以相互参考, 引导检索者从一个主题词找到另一个主题词, 以扩大或缩小选词范围, 从而达到检索全面、准确的目的。如: ① 疾病与其原因: Factor XIII Deficiency — Factor XIIIa; ② 器官与生理过程: Bone and Bones — Osteogenesis; ③ 器官与药物对它的作用: Bronchi — Bronchoconstrictor Agents。相当于“参见参照”(see related 与 XR)。