



博
学



JICHU YIXUE XILIE

基础医学系列

医学信息检索与利用

● 主编 夏知平

(第三版)

复旦博学·基础医学系列

复旦博学·基础医学系列

复旦博学·基础医学系列

复旦大学出版社



基础医学系列

医学信息检索与利用

主 编 夏知平

(第三版)

编写者 (以姓氏笔画为序)

王宇芳	李晓玲	俞 健
莫梅琦	夏知平	符礼平

復旦大學出版社

图书在版编目(CIP)数据

医学信息检索与利用/夏知平主编.—3版.—上海:
复旦大学出版社,2004.3

ISBN 7-309-03877-0

I. 医… II. 夏… III. 医药学-情报检索
IV. G252.7

中国版本图书馆 CIP 数据核字(2004)第 003157 号

医学信息检索与利用(第三版)

夏知平 主编

出版发行 **复旦大学出版社**

上海市国权路 579 号 邮编 200433

86-21-65118853(发行部) 86-21-65109143(邮购)

fupnet@fudanpress.com <http://www.fudanpress.com>

责任编辑 傅淑娟

装帧设计 马晓霞

总编辑 高若海

出品人 贺圣遂

印刷 崇明裕安印刷厂

开本 787×1092 1/16

印张 17.25 插页 2

字数 409 千

版次 2004 年 3 月第三版 2004 年 3 月第一次印刷

印数 1—3 100

书号 ISBN 7-309-03877-0/R·831

定价 30.00 元

如有印装质量问题, 请向复旦大学出版社发行部调换。

版权所有 侵权必究

主编简介

夏知平，男，1956年2月出生，华东师范大学图书馆学情报学专业本科毕业。现为复旦大学图书馆参考咨询部主任，副研究员。从事医学文献检索课教学19年，发表论文15篇，参编教材10本（包括全国统编教材2本）。1995年获上海医科大学优秀教育工作者称号，1996年获宝钢优秀教师奖，1997年获上海市教学成果（集体）二等奖。1998年7月至1999年7月在美国哥伦比亚大学医学图书馆作访问学者。

内 容 提 要

网络时代，信息浩瀚。《医学信息检索与利用》提供了快速有效获取医学信息的技能与方法。根据医药卫生工作者的信息需求，本书介绍了信息检索基础知识、图书馆资源利用、原始文献获取、常用生物医学文摘数据库、全文数据库、引文检索、Web of Knowledge学术信息资源体系、搜索引擎、常用医学网站、网上免费电子期刊查询、在线医学图谱、生物信息数据库、药物信息检索、专利检索、会议论文检索、学位论文检索、循证医学信息检查、医学信息调查研究、医学综述文献写作等内容。

依托复旦大学丰富的数据库资源，全书基本反映了国内外现有的与生物医学有关的重要数据库。本书内容全面，系统实用，既是医药卫生专业研究生和本科生的教材，又可作为医药卫生工作者继续教育用书和案头参考书。

前 言

医学信息检索与利用课程是一门实践性强、应用性广、内容更新快的科学方法课。在高校开设此课旨在培养学生敏锐的信息意识,学会探索知识的本领,掌握信息检索的技能,具备信息利用能力和独立研究能力。

1984年我国教育部发出关于在高等院校开设“文献检索与利用”课的文件以来,全国各类高校陆续为研究生、本科生开设了“文献检索与利用”课。信息技术的发展和网络应用的普及,使医药院校中医学信息检索与利用课程的教学内容不断丰富,教学方法趋向规范,授课面也越来越广。

近年来,文献数据库品种不断增加,因特网上的医学资源越加丰富,出版一本反映最新医学信息检索与利用知识的教科书迫在眉睫。复旦大学图书馆的6位教师在半年的时间里为此作出了不懈的努力。

《医学信息检索与利用》(第3版)在李晓玲主编的第2版的基础上,教材内容作了一些调整。充实与增加的内容主要有:图书馆资源利用、中外文全文数据库、Web of Knowledge学术信息资源体系、网络检索工具、网络专类信息检索等。删减压缩的内容有:书本式检索工具、医学写作等。

新版教材共10章63节,章节中有检索举例,章节之间有参见,章节后有思考题。第一章“信息检索基础”突出概念、分类、原理、规则、策略、技术,是学习本课程的基础知识。第二章“图书馆资源利用”内容简洁明了,全面实用,若课时有限,可由学生自学。第三~第六章为数据库检索,是本课程的重点学习部分。Medline与PubMed因其权威和使用普遍,用了较大篇幅予以介绍。依托复旦大学丰富的数据库资源,本书共介绍了11种中外文全文数据库。引文数据库因其特有的用途,单列出“引文检索”一章讲解。“ISI Web of Knowledge”国内目前完整订购的只有少数几所大学,因此教学中对本章内容可根据客观条件进行调整。第七~第九章为因特网信息检索与介绍。第七章中的Yahoo Search和Google知名度高,实用性强,应为必修内容。第八章“常用医学网站”和第九章“网络专类信息检索”内容丰富,建议对这两章内容泛讲,或择其部分内容细讲,留出一定课时,让学生按自己的专业和爱好,自主选择内容深入学习和上机操练。第十章“医学信息调查研究与综述写作”中,编写者对医学信息的搜集整理、医学信息的分析方法、医学调研报告和综述文献的写作,作了生动的举例和精辟的阐述。

本教材的使用对象是医药卫生专业的研究生、七年制学生和五年制本科生,也可用于医药院校教师、医师、科技人员继续教育的教材。

新版教材的编写得到了复旦大学图书馆徐一新副馆长及其他馆领导的支持,在此深表感谢。

由于网络信息变化更新快,因此在使用本教材时一定会发现某些数据库或网页的界面已与实际不符。同时由于编者水平有限和编写时间紧迫,教材中错误疏漏难免,恳请读者批评指正。

编者

2004年2月

目 录

第一章 信息检索基础	(1)
第一节 信息与信息类型.....	(1)
第二节 计算机检索与数据库.....	(6)
第三节 光盘存储技术与网络代理服务器.....	(12)
第四节 检索式编写与检索策略调整.....	(16)
第五节 信息素养能力培养与课程学习目标.....	(20)
第二章 图书馆资源利用	(23)
第一节 馆藏书刊排架.....	(23)
第二节 馆藏目录查询.....	(26)
第三节 图书馆服务利用.....	(27)
第四节 医学参考工具书.....	(30)
第五节 原始文献获取途径.....	(34)
第三章 文摘数据库	(36)
第一节 Medline	(36)
第二节 PubMed	(51)
第三节 中国生物医学文献数据库(CBMdisc)	(59)
第四节 EMBase	(65)
第五节 美国化学文摘.....	(70)
第四章 全文数据库	(77)
第一节 中国期刊全文数据库.....	(77)
第二节 中文科技期刊数据库.....	(81)
第三节 万方数据资源系统.....	(87)
第四节 超星数字图书馆.....	(92)
第五节 Elsevier Science(SDOS)全文数据库	(96)
第六节 EBSCO	(100)
第七节 ProQuest Medical Library	(105)
第八节 OVID 全文期刊库	(109)
第九节 OCLC FirstSearch	(113)
第十节 Kluwer 全文数据库	(117)
第十一节 Springer LINK 电子期刊服务系统	(119)
第十二节 全文数据库检索小结.....	(123)
第五章 引文检索	(126)

第一节	引文检索概述	(126)
第二节	SCI 简介	(127)
第三节	Web of Science	(129)
第四节	中文引文检索	(133)
第五节	期刊引用报告	(140)
第六章	ISI Web of Knowledge 学术信息资源体系	(142)
第一节	Web of Knowledge 概述	(142)
第二节	ISI Proceedings	(147)
第三节	DERWENT Innovations Index	(149)
第四节	BIOSIS Previews	(152)
第五节	ISI Chemistry	(153)
第六节	ISI Current Contents Connect	(156)
第七节	ISI Journal Citation Reports	(159)
第七章	网络检索工具	(162)
第一节	网络检索工具概述	(162)
第二节	综合型网络检索工具	(164)
第三节	医学网络检索工具	(171)
第八章	常用医学网站	(179)
第一节	国家科技图书文献中心	(179)
第二节	中国高等教育文献保障系统	(185)
第三节	常用中文医药卫生网站	(187)
第四节	美国国立卫生研究院	(189)
第五节	美国国立癌症研究所	(192)
第六节	美国国立医学图书馆	(194)
第七节	世界卫生组织	(200)
第八节	国外其他常用医药卫生网站	(202)
第九章	网络专类信息检索	(203)
第一节	机构与个人信息查询	(203)
第二节	免费电子期刊查询	(204)
第三节	在线医学图谱	(209)
第四节	药物信息查询	(213)
第五节	医学会议信息检索	(219)
第六节	学位论文检索及全文获取	(224)
第七节	生物信息数据库	(230)
第八节	专利检索	(238)
第九节	循证医学信息检索	(241)
第十章	医学信息调查研究与综述写作	(250)
第一节	医学信息调查研究概述	(250)

第二节	医学信息调查与研究资料的搜集、鉴别与整理	(253)
第三节	医学信息分析方法	(257)
第四节	信息调研报告及综述的写作	(262)
参考文献	(268)

第一章 信息检索基础

第一节 信息与信息类型

一、信息(information)与情报、文献

对 information 这一词的解释有几十种。作为比较正式的学术名称,主要有两种译名:信息、情报,台湾学者译为资讯。

关于信息的概念,不同的专业领域如图书情报学领域、计算机与通信科学领域等都有不同角度的解释。英国科学家波普尔(K. Popper)认为信息的概念可以分成三大类。第一类是有关客观物质世界的信息,即信息是事物存在方式及其运动规律、特点的外在表现形式。第二类是有关人类主观精神世界的信息,它反映人类所感受的事物运动状态及其变化方式,处于意识和思维状态的信息。第三类是有关概念世界的信息,它反映人类所表述的事物运动状态及其变化方式,用语言、文字、图像、影视数据等各种载体来表示。

信息、文献与情报既有联系又有区分,他们都与知识相沟通。

信息是知识的原料,知识是信息的产品。知识是信息的一部分。人类在认识世界和改造世界的过程中,不断接受客观事物发出的信息,经过思维加工,获得了对事物本质及其运动规律的认识,信息如此转化为知识。人类获取知识以后,再将这些知识用来创造新信息,获取新知识。如此反复循环,信息越来越纷繁,知识越来越丰富,知识不断提高和深化。因此,人类要认识世界和改造世界就必须不断地搜集信息、加工信息、创造信息,使信息造福于人类。

国际经济合作与发展组织(OECD)在 1996 年发表的《以知识为基础的经济》报告中系统地提出了知识的 4 个 W 概念:知道是什么(know-what);知道为什么(know-why);知道怎么做(know-how);知道是谁(know-who)。这是目前人们对知识的一般认识和概念划分。

当人们为了解决某一个特定问题去搜寻所需要的知识,那一部分知识就是情报,其含义国际标准定为:被传递的知识或事实。钱学森称它是激活了、活化了的的知识。情报来源于知识,必须在特定的时间里经过传递,并能为用户所接受和利用。情报作为交流对象的有用的知识,是知识的一部分,是进入人类社会交流系统的运动着的知识。这些含义包括了情报的 3 个基本属性——知识性、传递性、效用性。

1. 知识性

从情报角度来讲,情报来源于知识,而知识又来源于信息。任何情报都具有一定的知识和信息,但并非所有的知识、信息都能构成情报,只有那些经过加工并为用户所需要的特定的知识或信息,才称得上情报。

2. 传递性

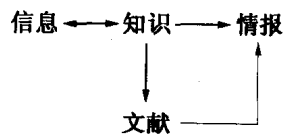
知识、信息要转化为情报,必须经过传递,并为用户接受和利用。通常记录在书刊中的知识属于静态的知识,还没有为人所用。只有当书刊中的知识传递给了用户,并发挥了使用的价值,才使静态的知识变为动态的情报。

3. 效用性

效用性是衡量情报服务工作优劣的重要标识。人们创造情报、传递情报的目的就在于充分利用,提高其效用性。情报的效用性表现为启迪思维,增进见识,改变知识结构,提高认识能力,帮助人们改造世界,发挥其使用价值、社会价值和经济价值。在知识经济社会中,人们也把一部分能够增值的情报称为竞争性情报(competitive intelligence)。

关于文献(literature, document),我国国家标准《文献著录总则》中明确定义为:“文献是记载有知识的一切载体。”文献是知识的外在表现形式。文献中记录着大量的知识和信息,这些知识和信息为读者所利用时就转化成了情报。长期以来,人们已经习惯从文献中获取情报,把它看成是一种重要的情报源。但文献并不是唯一的情报源,除了文献情报源还有非文献情报源。非文献情报源主要指的是实物情报和口头情报源。

由此信息、文献、情报三者之间的关系就可以写成这样的关系式:



医学科研成果大多数是以文献的形式加以记载并得到学术认可。医学文献是提供医学科研和临床实践借鉴、交流的重要载体。医学科研人员在科研过程中,包括立项、试验研究、成果鉴定、交流推广,都要通过文献查阅、信息调研来了解课题的相关信息,如该课题是否有人进行过研究、研究的程度如何、已经有哪些重要的突破、有哪些问题有待解决、研究发展的趋势如何等等。医学信息包括知识、文献和情报是促进推动医学科研发展的重要支撑和资源动力。

二、信息的类型

信息按照其外在的表现形式和内涵等,有许多不同的分类法。了解信息的类型,对获取信息、分析信息、表达信息都很有必要。

1. 网络环境下信息资源

(1) 网络环境

20世纪90年代初,以美国为首的一些发达国家率先开发、研究建立国家信息基础结构(National Information Infrastructure, NII),即建立遍布全国的由通信网、数据库计算机以及电子产品组成一个完备的信息网络。全世界各国包括我国都在着手进行国家信息基础结构NII的建设工作。在NII中采用各种现代信息技术。例如,由互联网络(包括因特网)、光缆、宽带传输、微波、卫星系统等技术组成的通信系统,采用高性能、高效能、便携式以及带手写语音输入等各种新一代的计算机,建立各种公用和专用数据库、带视频/音频/图像的数字资料库等等。

由此而带来的集计算机技术、通信技术、激光技术、自动控制技术、光导技术和人工智能

技术等之大成的信息技术突飞猛进,使得图书馆、情报服务技术发生了翻天覆地的变化。计算机技术、声像技术、复制技术、网络技术、光盘技术、全文存储技术和多媒体技术等已成为图书馆信息服务工作的技术基础。数字图书馆研究、自动化网络研究、自动语言规范技术、检索方法学、知识库的建设、智能检索系统、自然语言处理等技术正成为图书馆信息服务工作的重要研究课题。

医学信息的传递形式由一般的信件、会议、人与人口头的交流拓宽到电子信件、可视会议、远程医疗、远程科研。人们可以在计算机屏幕前,瞬时地交流医学信息,包括从科研背景、人物介绍、相关文献到实时影像如 MR、CT,通过计算机网络探讨治疗诊断问题等。通过信息基础结构高速传输,作为信息存储和利用的机构——图书馆可以不仅仅是一个实体,医学信息资源在全球范围的一定程度上已达到共享。

(2) 电子信息资源

电子信息资源主要有单行版电子出版物(光盘、磁带等)和网络信息资源两大类。

随着计算机及网络技术的发展,电子信息计算机阅读型的文献越来越受到人们的青睐。它以数字形式将文献存储在磁、光等介质(如磁盘、光盘等)——计算机的外存储中,通过本地计算机或远程通信传输的网络计算机系统来进行阅读,其中有文本(text)型、多媒体(multimedia)型、超文本(hypertext)型等。

文本型信息是电子医学信息、知识获取的主要热点。随着数字图书馆、全文数据库的涌现,在线阅读、获取医学信息与知识已成为医学科研临床工作者的热切愿望。

多媒体文献是以数字形式包括计算机视频、音频等技术处理来记载图形、影像、动画以及实时动态信息,并让用户既可通过计算机阅读,又可与计算机进行互相问答所谓交互式的信息利用形式。医学科研人员、临床医师在计算机屏幕前获取医学信息从基础知识(包括人体解剖多媒体逐层分析、病理切片、化学物质空间构架)到临床实践(外科手术、心脏听诊模拟等);从书目文献到生物信息基因序列乃至知识库的获取,无不体验到现代信息技术给医学信息带来的无穷魅力。

超文本(包括超媒体)文献,则是由于计算机网络特别是 WWW 技术的发展而产生的一种通过文本或图像的关键词或图标链接文件的形式,让用户不必考虑信息的来源或分类,随意地在网络中寻找相关信息、文献的一种新型计算机信息阅读的方式。

支持文献阅读的软件,现在有许多种,如在网上比较普遍的有 PDF 格式(用 Acrobat reader 软件阅读)、超文本标记语言 HTML(Hypertext Mark Language)、可扩展标记语言 XHTML(eXtensible Hypertext Markup Language)、CAJ(中国期刊网阅读软件)等等。它们支持着不同数据库文献的阅读、保存和打印。在进行不同数据库文献阅读时,都需要根据特定数据库的要求来下载不同的阅读软件。

电子信息资源中最具代表性的事物便是数字图书馆。对于数字图书馆,目前尚未有明确的定义,但人们已经有了一些初步的概念,如有人认为它是“全球信息高速公路上信息资源的基本组织形式”,有人则认为它是“存储电子格式的资料,并对这些资料进行有效的操作”。不管定义如何,数字图书馆都有以下一些特点:信息存取多媒体化、信息组织有序化、操作电脑化、传输远程网络化、资源共享化和结构连接化(跨库连接无缝化)。

数字图书馆从资料保存的安全性和使用的方便角度上来看,可以分成以下几个级别。

永久保存级:具有保存价值的资料存储在图书馆硬盘或高等存储装置上。这种保存方

式,资料数据不容易丢失。局域网上数据的使用,也比较稳定。

服务级:通过网络传输,利用其他图书馆计算机服务器上的资源,资料在本地图书馆不具有永久保存性。通常资源使用需要专线,用户使用有权限控制。

镜像级:其他数字图书馆资料拷贝到本图书馆网站,保存责任在其他图书馆,使用权受网络 IP 地址控制,用户使用有一定的范围限制。

链接级:通过网站的链接,不受控制地获取和保存,但由于网址的变迁、网络传输、数据源的变化,链接容易失效。

(3) 传统文献(印刷型、缩微型、视听型)资源

1) **印刷型:**从过去的手工书写发展到由计算机打印,或复印机复制、网络传递等形式。不管记录方式如何,载体总是纸张。由于印刷型文献符合人类传统的阅读习惯,且实用、方便,将在相当长的时间内与电子型文献共存。

2) **缩微型:**是指采用照相技术,将文献以 1:100 或 1:1000 的比例,缩小存储在感光胶卷或平片上,通过专门的阅读机进行阅读。其特点是,存储容量较印刷型大,保存期比印刷型长。

3) **声像型:**包括录音带、录像带、幻灯片、电影拷贝,目前多数以音频或视频的电子形式在光盘和计算机网络中出现。

2. 按文献发布的类型分类

(1) 图书(book)

图书是指一些记录的知识比较系统、成熟的文献。医学科研和临床所涉及的图书有教科书、丛书、专著、全集、会议论文集以及词典、百科全书、指南、手册等参考工具书等。

教科书是供医学生和医学工作者进行专业学习的主要医学文献。大多是某个专业的研究总结,反映了较成熟的医学理论,具有严密的系统性和逻辑性,内容可靠性强。在科研课题资料查找时,可以从教科书开始,明确专业基本概念和理论。教科书所附的参考文献选录经典,质量也较高。

专著是以一个专题为中心的科学著作,如《休克》、《心血管药理学》等。专著对某一个专题有较深入的研究知识和独到见解,是查阅某一个课题的“第二阶梯”资料。在阅读了教科书后,紧接着要深入了解专题内容知识,就应读专著。

参考工具书是供日常工作、阅读或写作中随时查阅用的一类文献,它内容有序,便于查考。通常也在某一课题开始时,用于搞清一些名词、术语、数据等反映知识点的资料。医学生和医学工作者可以常备综合性的词典解决文字的规范问题,辞海和美国的 *Dictionary of Dorlen* 则兼有百科全书功用,也是必备的工具。为准确使用医学名词术语,还要有一本医学词汇或词典。

(2) 期刊(journal)

期刊是指一些记录的知识比较新颖、所含信息密度比较大的连续出版物,一般都有固定的名称。如:有关自然科学的综合性期刊有《中国科学》、《自然杂志》等;医学学科的专业性期刊《中华医学杂志》、《中国药理学学报》等;学报如《复旦大学学报》;检索性期刊如《美国化学文摘》(*Chemical Abstracts, CA*)、《美国生物学文摘》(*Biological Abstracts, BA*)、《美国医学索引》(*Index Medicus*)等。期刊通常刊登了能够反映学科领域最新的理论、方法、技术等论文(*journal article*)、综述(*review*)、病例报告(*case report*)等各种文献信息。

期刊论文包括研究报告、论著、著述等是反映科研最新成果的科学论述文献,是科学研究原始创造的首次记录。它们是一类具有科学性、学术性、创新性特点的医学文献,是医学科研工作者在科研课题进行的全过程中,包括立题、试验研究、成果鉴定都要及时查阅的文献。

综述文献是综合描述某一专题或学科在一定时间内研究的现状和进展的文献。该类文献综合性强且有较高的权威性,能够直接反映专业领域科研的动向和情况,也是医学科研人员在课题开始进行时,为了了解科研背景、现状、预测前景,而要阅读的医学文献。

期刊是科研课题工作的主要文献源、信息源、情报源。医学科研工作者依靠期刊来及时跟踪最新的国际、国内的研究动向,把握科研的主动权。

(3) 特种文献(special literature)

特种文献是指无法归入图书或期刊的文献,比如科技报告、学位论文、专利说明书、标准文献、会议文献、政府报告等。这些文献一般不公开出版,普通图书馆也不收藏。但是,特种文献反映的有许多涉及了最新的研究和技术以及国家的法规、标准定义等信息,也是医学科研的重要信息源。

3. 文献按揭示信息内容程度不同的分类

(1) 一次文献

也指原始文献(primary literature),是指记录原始的创造和科研成果,即记录新理论、新技术、新知识、新发明、新见解的一类文献。如期刊论文、学位论文、专利文献、会议文献等,一次文献目前主要有以下特点。

数量激增,种类繁多:有资料报道,近 20 年的科学文献的量相当于人类历史几千年来总和,并以每 10 年翻一番的速度增长。随着社会的发展,医学信息的产出数量呈指数级增长。有资料报道,现在医学文献每天发表 12 000 多篇。知识学科内容彼此相融、交叉,分支学科、边缘学科大量涌现。17 世纪,医学期刊只有 10 种;20 世纪初,医学期刊有 1 600 多种;20 世纪末医学期刊有 21 000 多种;目前有 35 000 多种,占科技期刊的 1/5。另外,目前在全世界发行的医学文献有几十个语种。

发表分散,老化加快:学科的分支越来越细,造成原文发表分散,如关于免疫学的文献可以出现在肿瘤学、分子生物学、医学工程等多种杂志上。科学在发展,知识要更新,文献就会老化,科技文献的寿命一般为 5~10 年,而医学文献的老化速度更快,其“半衰期”一般为 5 年。

数字化与印刷型并存:目前,全文数据库大量涌现,各种出版商以及数据库公司合作开发的全文数据库不下几十种。通过网络和光盘等载体,提供给用户及图书馆近万种电子期刊,用户在数字图书馆中可以更方便的检索全文,直接获取一次文献。但是,由于知识产权和版权等问題,目前有一些重要、权威期刊尚未收录入全文数据库,而阅读这些期刊中的文献只能通过印刷型载体来进行。

另外,全文数据库由于主建单位使用的检索平台、软件各异,检索使用的界面、语言、句法、规则的多样化,给用户使用平添了诸多麻烦。由此,全文数据库的异构平台正在悄然兴起。在统一的检索平台下,用户使用一种检索提问,将可在多个数据库中进行同时、有效的检索。

(2) 二次文献

主要指书目类(bibliography)文献,包括题录型和文摘型文献。目前都以计算机网络化的数据库形式加以处理和应用,是对原始信息进行加工整理组织后,便于管理和查找利用原始信息的工具。书目文献是将文献的概要内容,以及文献中的题名、著者、主题、原文的出处

(刊登的期刊名称、年、卷期页、网址等)、收藏文献的图书馆或机构等加以记录,并按一定规律和方法编制成的检索工具。此类文献常见来源有:图书馆馆藏目录和书目文献数据库。

书目文献数据库主要有题录和文摘两种著录格式。

题录:通常只提供文献的著者、题名、原文的出处(原文登载的期刊名、发表的年、卷期、页)等简要信息,是获取全文的线索。下例是中国生物医学文献数据库的题录格式。

标 题:蜂毒肽对细胞膜跨膜离子转运的作用

著 者:杨申; Gaspar CARRASQUER

出 处:中国药理学报 1997.01.15; 18(1): 3-5

文摘:随着计算机检索技术和情报学技术的发展,包括计算机的处理能力和存储容量不断提高和扩大,数据库文献进一步发展到文摘型著录。文摘型文献既提供了文献的主要内容、主题梗概,也是查找全文的线索,所以也具有检索性。20世纪70年代以前所指的文献检索,大部分指的就是这些类型的文本获取。由于二次文献数据库历史悠久,系统成熟,收录的期刊较全。目前,如要满足文献的查全,此种类型的数据库仍是首选工具。

(3) 三次文献

三次文献是在阅读一次文献的基础上,分析综合归纳信息后,组织形成具有资料性、查考性、阅读性的文献,如教科书、综述、参考工具书、进展、调查报告等。

参考工具书通常包括了数值数据(data)和事实(fact):如,实验室各种数据、仪器的参数、图表、化学物质的理化常数等。反映信息内在含义、知识点的内容,回答某些特定的事实和具体的问题。如什么是克隆技术?阿司匹林的系统化学物质名称?等等。但是,随着信息资源的发展日益迅猛,许多新的知识已经不能仅仅依靠编纂滞后的参考工具书(词典、百科全书、年鉴、手册等)来获取,此时,网络搜索引擎则成为重要的参考工具。通过引擎的搜索,特别能够解决新的、跨专业的知识点信息,如查询“flanking 在遗传学研究中的正确含义”一类信息。

(李晓玲)

第二节 计算机检索与数据库

计算机检索简称机检,是指利用计算机从文献数据库或因特网上查寻所需文献信息的过程。机检包括光盘检索、联机检索和网络检索。与手工检索相比,机检具有信息容量大、检索速度快、检索途径多、检索灵活、检索结果存储传递便捷等优点。

一、计算机检索发展简史

1954年,美国海军兵器实验所图书馆用IBM701型电子管计算机建立了世界上第一个情报检索系统以来,计算机检索的发展经历了以下几个阶段。

1. 脱机检索(off-line retrieval)

20世纪50年代问世的计算机检索是没有通讯线路支撑的单机检索,故称之为脱机检索。情报用户通过信函提交检索提问单,检索系统的专职检索员将积累到一定数量的检索提问输入计算机,由计算机批量进行处理,然后将打印结果寄给用户。1964年,美国国立医学图书馆研制成功的MEDLARS(Medical Literature Analysis and Retrieval System,医学文献分析与检索系统)就是当时的一个医学文献脱机检索系统。脱机检索固然比手工检索

先进,但存在着得到检索结果速度慢、情报用户无法与检索员对话的缺点。

2. 联机检索(on-line retrieval)

随着数字通讯技术的发展和分组交换公用数据通讯网的普及,20世纪60年代末、70年代初出现了联机检索。联机检索是计算机技术、通讯技术和现代情报技术三者结合的产物。联机检索时,用户的提问和联机数据库中的反馈数据借助通信线路来回传输。1971年,美国国内90个医学机构与美国国立医学图书馆的Medline(MEDLARS on-line)数据库联机。卫星通讯技术的发展推动了国际之间的联机检索服务,世界上相继出现了DIALOG、ORBIT等一些国际联机检索系统。即时检索远地文献数据,当场打印联机检索结果,是联机检索的一大优点。但是,联机检索收费高,尤其是昂贵的国际联机通讯费,阻碍了联机检索的广泛应用。

3. 光盘检索(optical disc retrieval)

光盘是一种用激光记录和读取信息的存储介质,具有容量大、成本低、寿命长、用途广、读取快等优点。尤其是光盘检索收费比联机检索低,因此1984年光盘问世后,光盘检索迅速普及。用户通过租赁或购买获得光盘数据库,再为计算机配上光盘驱动器和检索软件,就可以实地实时进行检索。1986年起我国开始引进Medline光盘,1992年我国第一个光盘数据库《中国科技期刊篇名数据库》研制成功。

光盘中最常见的是CD-ROM(Compact Disc Read Only Memory,高密度只读光盘)。CD-ROM只能从中读取数据,而不能写入数据。一张CD-ROM盘片上能够存储650 MB字节,相当于400多张1.4兆软盘的存储容量。其他类型的光盘有WORM(Write Once Read Many,一次写入光盘)和Rewritable(可擦写光盘)。

同一种数据库可以有不同的光盘版本,其操作方法和检索界面互不相同。例如Medline数据库的光盘产品有美国银盘公司(SilverPlatter)的WinSPIRS版,有OVID公司的OVID版,有DIALOG公司的DIALOG版等。

光盘检索以收费低廉吸引情报用户,但在数据库品种选择面上和数据更新速度上不及联机检索。因此,光盘数据库问世后,光盘检索与联机检索并存,两者各自发挥着自己的优势。

4. 网络检索

20世纪90年代进入了网络信息时代。一方面,光盘检索从单驱检索发展到塔网检索,局域网内的多台计算机工作站可同时检索塔式驱动器中的多张光盘。目前塔式驱动器又被大容量硬盘组或硬盘阵列所取代,多张光盘的数据内容可以镜像或缓存到局域网主机的硬盘上,大大提高了网络检索速度和共享能力。另一方面,因特网和World Wide Web的崛起,实现了全球范围的信息资源共享。浩如烟海的网上信息、易学易用的Web浏览器、查询便利的搜索引擎、不断增加的网络数据库,正在影响和改变着人们的工作、学习乃至生活方式。计算机信息检索与利用由此进入了一个崭新的历史时期。

二、数据库的类型与结构

机检的对象是数据库和因特网上的各种信息资源。数据库是指由计算机进行处理的一定数量同类信息的有序集合,是用来存储和查找文献信息的电子化检索工具。

1. 数据库的类型

(1) 书目数据库(bibliographic database)