

高等医药院校教材



医学信息学

主编：梁玲芳
蒋海萍
林 红
何立芳
朱象喜

中国档案出版社

责任编辑/刘 琛

封面设计/沈钰浩

图书在版编目(CIP)数据

医学信息学/梁玲芳等主编. —北京:中国档案出版社,
2006. 1

ISBN 7-80166-654-2

I. 医... II. 梁... III. 医学:信息学 IV. R-05

中国版本图书馆 CIP 数据核字(2005)第 160145 号

医学信息学

出版/中国档案出版社(北京市宣武区永安路 106 号)

发行/中国档案出版社

印刷/杭州萧山日报印务有限公司

规格/787×1092 1/16 印张/20 字数/480 千

版次/2006 年 1 月第 1 版 2006 年 1 月第 1 次印刷

印数/01—3000 册

定价/30.00 元

前 言

21 世纪是信息时代,科技知识的发展完全可以用“瞬息万变”来形容。当代大学生在掌握所学专业专业知识的同时,还必须掌握获取与时俱进的新的专业知识的手段和技能,只有这样,才能创造出辉煌的业绩,为医疗卫生事业做出贡献。

如果说在校的专业老师只能授以“鱼”(专业知识),那么信息素质的培养可谓授之以“渔”(获取新知识的技能)。本书强调对学生信息素质的培养,并提出培养的方法。学生在掌握了医学信息学的基本知识和查询技能之后,将终身受用,不断获取全世界新的专业知识,永不枯竭。

网络信息资源极其丰富,人们常以“海量”来形容——取之不尽,用之不竭。可是在这无穷无尽的信息海洋里鱼龙虾蟹混杂,其中医学文献数量极大,居各类科技文献之首,约占 20%~25%,增长速度也居领先地位,所以要想在这么浩渺的信息海洋里找出需要的医学信息,真好比“大海捞针”。

医学信息学就是专门研究医学信息资源有哪些类型,什么地方有哪些医学信息,怎样找到需要的医学信息,怎样获取这些信息,并如何利用这些信息用于科研医疗和撰写医学论文的一门新的学科。它是信息时代医学专业学生、医学科研人员、临床医务人员必须具备的基础知识。

本书内容包括:信息素质的培养,医学信息及查询的基础知识,网络信息查询的基础知识,校园网 13 种国内外著名全文数据库,国内外 8 种免费医学全文数据库,引文信息及影响因子查找方法,以及特种医学信息,包括循证医学、专利、学术会议、学位论文、药学文献、医药图谱、国内外免费医学继续教育课程、医院信息系统,国内外重要医药卫生机构网站简介,期刊的利用、医学情报调查研究、医学科技查新、医学科研设计、医学论文撰写与投稿技巧,等等。

本书具有下列特点:

1. 提出了信息素质的概念和内涵及其培养的方法。
2. 对重要术语的概念都给予明确的解释,并对“检索方式”、“导航栏”等术语提出特定的概念,避免了认识模糊、叫法不一等问题。
3. 本书列举出 14 种典型的检索式和调整检索策略的方法。
4. 借助浙江大学这个全国重点大学信息资源丰富这一平台,本书包含了 13 章由学校包库的国内外著名大型全文数据库。
5. 考虑到医学院校的学生从母校毕业后绝大多数走向各级医院和科研机构,他们很少有大学里包库的检索条件,为此本书专设 8 章详细介绍免费外文医学网站和电子医学期刊全文数据库及其查询方式。
6. 本书介绍的免费药文学文献信息查询、免费医药图谱查询、免费继续教育课程查询,以及医院信息系统、中国医院数字图书馆、被 Medline 和 SCI 收录的中国医学期刊和影响因子及其查找方法等内容更是同类书中所未见的。
7. 本书列出了国内外重要医药卫生管理、研究机构的网址,并作了简单介绍,便于访问。
8. 邀请专门从事有关专业研究的专家撰写他们熟悉的章节内容,如邀请中国医学科学院医学信息研究所专门从事医学信息研究数十年的老教授撰写“医学情报调查研究”,邀请在美国卫生研究院(NIH)从事生物信息工作的华人博士撰写“生物信息学数据库及其利用”,皆是理论与实际经验结合。

本书承浙江大学研究馆员、全国高等医学院校医学文献检索教学研究会顾问(原常务理事、教材研究组组长)朱象喜老师共同策划和负责总审,谨表谢意。

本书在编撰过程中因时间仓促,难免有差错,敬请同行老师、同学、读者指正赐教。

编 者

2005 年 10 月 30 日于杭州

目 录

前言	1
第一篇 医学信息学总论	1
第一章 《医学信息学》教学的目的意义	1
第一节 《医学信息学》课程的发展历程	1
第二节 《医学信息学》教学的目的意义	1
第二章 医学信息学基础知识	3
第一节 信息和文献	3
第二节 信息检索和文献检索	5
第三节 书本式医学文献检索工具	8
第四节 数据库的结构和类型	10
第五节 网络信息资源	11
第三章 参考工具书	12
第一节 参考工具书的定义与特性	12
第二节 参考工具书的排检方法	13
第三节 参考工具书的类型与医学参考工具书举要	14
第四节 医学参考工具书的利用	15
第二篇 网络信息检索	19
第一章 网络信息检索基础知识	19
第一节 Internet 概述	19
第二节 获取网络信息资源的主要工具	22
第三节 网络信息检索资源服务器	24
第四节 检索途径	24
第五节 检索方式	26
第六节 网络信息检索技术	26
第七节 检索方法和步骤	29
第八节 制订检索策略、编写检索式和检索策略的调整	30
第二章 国内著名综合搜索引擎	32
第一节 百度	33
第二节 搜狐	36
第三节 新浪	36
第三章 国外著名综合搜索引擎	36
第一节 Google	36
第二节 Yahoo!	39
第三节 ProFusion	42

第四节	Dogpile	45
第五节	Metacrawler	47
第四章	国内著名医学搜索引擎	49
第一节	39 健康网	49
第二节	37℃ 医学网	51
第三节	导医网	52
第五章	国外著名医学搜索引擎	52
第一节	Medscape	52
第二节	Medhunt 与 HONselect	55
第三节	Healthweb	58
第四节	Oncolink	59
第五节	Medical Matrix	61
第三篇	校园网数据库	63
第一章	中国生物医学文献数据	63
第二章	维普信息资源系统	72
第一节	《中文科技期刊数据库》(全文版)	72
第二节	《中文科技期刊数据库(引文版)》	81
第三节	维普医药信息资源系统	83
第三章	清华万方数据库检索系统	88
第一节	关于 CNKI 工程及其与清华同方的关系	88
第二节	《中国学术期刊全文数据库》(新版)	89
第三节	《中国期刊全文数据库》(旧版)	94
第四章	中国医院数字图书馆	100
第五章	万方数据资源系统	107
第一节	数据资源系统(Web 版)简介	107
第二节	数字化期刊子系统的检索方法	109
第六章	MEDLINE 光盘检索	118
第一节	《医学主题词表》	118
第二节	WinSPIRS 系统的 MEDLINE 检索	121
第七章	ProQuest Medical Library 全文数据库	131
第八章	OCLC FirstSearch	136
第九章	OVID 全文期刊数据库	139
第十章	EBSCOhost 数据库	145
第十一章	Elsevier Science(SDOS)	149
第十二章	Kluwer Online Journals	151
第十三章	SpringerLink 全文数据库	153
第四篇	免费全文医药文献信息	157
第一章	PubMed	157
第二章	Amedeo(Free Medical Journals)	165
第一节	FreeMedicalJournals.com	165
第二节	Amedeo.com	169

第三章	HighWire Press	172
第四章	日内瓦医学教育研究基金会.....	179
第五章	巴西 Sao Paulo 大学医学院图书馆	183
第六章	加拿大 McGill 大学卫生科学图书馆	186
第七章	Internet Drug News.com	191
第八章	生物信息学数据库及其利用.....	199
第五篇	引文信息	212
第一章	Web of Science	212
第二章	期刊引用报告(JCR)	218
第六篇	特种文献信息	225
第一章	循证医学证据信息.....	225
第二章	专利文献信息.....	236
第三章	医学学术会议文献信息.....	242
第四章	学位论文信息.....	246
第五章	药学信息.....	251
第六章	医学图谱信息.....	258
第七章	医学继续教育信息.....	263
第八章	医院信息系统(HIS)	266
第七篇	国内外重要医药组织机构信息	276
第八篇	文献信息的利用	280
第一章	电子期刊的利用.....	280
第二章	医学情报调查研究.....	285
第三章	医学科技查新.....	296
第四章	医学科研设计.....	299
第五章	医学论文的撰写和投稿.....	302
参考文献	312

第一篇 医学信息学总论

第一章 《医学信息学》教学的目的意义

医学信息学是专门研究医药信息的概念、范畴、类型、分布、检索和利用的一门新兴的学科。它是信息时代医学专业学生、医学科研人员、临床医务人员必须具备的基础知识。

信息检索被人们喻为开启知识宝库的钥匙。当你在学习、工作乃至生活保健方面遇到疑难,或要开展一项新的研究,或要完成一篇论文、一种发明……都可以借助信息检索,从一个陌生者入门,或从略知一二到获知甚多,即使已学富五车,也还需要进入到一个更丰富的知识宝库。人类社会已迈入知识经济时代,作为个人,谁掌握了知识与信息并用来创新,谁就掌握了生存与发展的主动权;作为国家,富国强兵的首要战略,就是争取信息与知识更新的优先权,从而转化为科技优势。知识经济时代对人才培养和高等教育提出了从应试教育向素质教育过渡和注重学生信息素质培养的要求。在医药高校和卫技系统,这一任务分派给医学信息检索课来承担。

第一节 《医学文献检索》课程的发展历程

教育部分别于1984年和1985年发布了《关于在高等学校开设〈文献检索与利用课〉的意见》和《关于改进和发展文献检索教学的几点意见》,目的在于“提高大学生的自学能力和独立研究能力”,把学生从一个知识型人才培养成为素质型人才,特别要注重学生自学和独立研究能力的培养。要求“凡是有条件的学校可作为必修课,不具备条件的学校可作为选修课或先开专题讲座,然后逐步发展完善”。这些文件的颁布,使文献检索课教学有了明确的发展方向,走向正规化。

在国家教委1987年颁发的《普通高等学校图书馆规程》中也明确规定:“高等学校图书馆应组织力量,采用多种方式对读者进行系统的检索与利用文献的教育,学校应将其列入教育计划。”1992年国家教委在总结我国高校文献检索教学工作的基础上,又颁布了《文献检索课教学基本要求》。

1999年的《中共中央国务院关于深化教育改革全面推进素质教育的决定》对高等教育实施素质教育提出了明确要求:“高等学校要重视培养大学生的创新能力、实践能力和创业精神,普遍提高大学生的人文素养和科学素质。”教育部2001年工作要点中提出坚持用最新的科学文化成果教育学生,融传授知识、培养能力和提高素质为一体,促使学生广泛参与科研和社会实践,加大对学生创新精神与实践能力的培养力度。

国家教育部2002年颁布的《普通高等学校图书馆规程(修订)》中再次把“开展信息素质教育,培养读者的信息意识和获取、利用文献信息的能力”作为高等院校图书馆基本任务之一。新的国家技术监督局与教育部等文件,认定文献检索学为三级学科。

第二节 《医学信息学》教学的目的意义

《医学信息学》这门课程,研究对象是医学文献与信息;研究的内容是医学文献信息资源的分布、类型、查检、传递和利用的方法;研究的目的是提高学生的信息意识及获取和运用文献资源的技能,提高学生自学能力和独立研究问题的能力,为科技创新提供知识。

一、信息素质

信息素质指对信息需求的迫切程度和获取、评价及利用信息的能力。信息素质的内含包括信息意识、信息能力和信息道德三个方面。

1. 信息意识。也称情报意识,指人们对信息重要性的认识程度,对信息需求的迫切程度,寻找信息的自觉程度,分析信息、判断信息和吸收信息的洞察程度。信息意识强的人认为,信息对人类社会、经济的发展,人才的培养成长,成果的获得等具有极其重要的作用,把搜集信息作为事业建树之道,对信息具有特殊的敏感性,会积极主动随时密切关注国内外新的信息,自觉地与自身知识更新、工作实践联系起来,启发创新思维,做出科学的决策,为教学、科研、医疗事业做出贡献;而信息意识淡薄者缺乏改革的信息思路,不善于发觉、捕捉、利用信息,无法及时跟踪学科发展前沿的动态,最后导致重复他人的科研项目,劳而无功,业绩平庸。

2. 信息能力。指对信息资源了解程度和掌握获取、分析和利用信息的方法和技能的程度。信息能力强的人了解信息资源甚深,能快速有效地获取需要的信息用于教学科研医疗。

3. 信息道德。指获取他人的信息时必须遵守法律法规(尤其指版权法和专利法)的道德品质。

二、信息素质的培养

(一) 信息意识的培养

信息,实际上也是知识,拥有某方面的信息就是拥有了某方面的知识。

1. 信息作用观念的培养。①首先要增强对信息作用的认识。“知己知彼,百战不殆”就是关于信息对于克敌制胜重要意义的概括。②增强信息是资产、是财富、是力量的认识。第二次世界大战后的日本,正是依靠信息为先导来富民强国;工程院院士郑树森 1995 年开始肝移植手术,得益于国外 1968 年以来大量人体肝移植的信息。③加强竞争意识,要培养坚强的意志,培养敢想敢干,勇于创新、创造的精神及勇于和敢于迎接挑战、挫折与艰辛的心理素质。④增强优胜劣汰的危机感的认识。缺乏信息头脑,没有创新能力,就会有被辞退的可能。只有经常不断地给自己“充电”,汲取新知识,善于开拓创新,才能立于不败之地。

2. 提高政治思想觉悟,学好本领,不断更新知识,开拓创新,树立为国争光,为社会人民作贡献,为卫生事业作贡献的责任感。

3. 树立千方百计采集情报信息为病人解除病痛和疑难病症的责任心。

(二) 信息能力的培养

1. 信息资源知识的培养。①要熟悉图书馆各种馆藏,充分利用图书馆,提高自己的知识水平。②要掌握网上各种不同的医学信息资源,汲取先进知识,为我所用。

2. 信息获取能力的培养。要熟练掌握光盘检索和网络信息检索知识和技能。

3. 信息分析和利用能力的培养。要学会对获得的信息进行分析,取其精华,去其糟粕,利用各种信息分析方法,找出当前某个课题的热点、冷点,寻找两个学科的交叉点作为研究的课题等。

(三) 信息道德的培养

科技的发展日新月异,人类的进步总是站在前人的肩膀上,有所发现、有所发明、有所创新、有所前进。前人的科研成果是宝贵的财富。前人的论文、专利、技术方法、经验法则可以学习、引用,作为自己研究的起点,但必须在法律法规允许的范围内使用,并注明文献的出处,绝不能贪他人之功,改头换面,占为己有。

第二章 医学信息学基础知识

第一节 信息和文献

一、信息、知识、情报与文献的概念及相互关系

1. 信息(Information) 比较普遍的认识是:信息是自然界、人类社会以及人类思维活动中普遍存在的,泛指一切事物(包括物质的或精神的)运动的状态及运动的方式。简而言之,是事物存在和运动的状态及其特征的反映。

不同事物具有不同的运动状态、特征和方式,因而产生不同的信息。通常,人们根据信息发生源的不同,将信息分为四大类:自然信息、生物信息、电子信息和社会信息。如:打雷、闪电等是自然信息,体温升降、鸟语花香等是生物信息,无线电波、计算机使用的代码和脉冲信号等是电子信息,人类的社会活动及所使用的语言、文字、图形符号等是社会信息。医学信息包括上述的四类信息,它们都是重要的资源,与材料、能源一起构成了现代社会发展的三大支柱。

2. 知识(Knowledge) 来源于信息,是人类对大量信息进行思维分析,加工提炼,并加以系统和深化而形成的结果。简而言之,理性化、优化和系统化了的的信息即为知识。信息与知识的关系,犹如原料与产品的关系。

知识按其内容可分为两大类:自然科学知识和社会科学知识。医学知识属于自然科学的范畴,是人们在与疾病作斗争的实践过程中积累起来的经验结晶。

3. 情报(Information) 情报最初用于战争,“战时敌情之报告”。后来应用领域越来越广,如经济情报、科技情报等。情报是针对一定的对象的需要而传递有参考价值的新信息和新的知识。简单地说是有用知识的传递,这里包括了情报的三个特性:效用性、知识性和传递性。例如,国外某科研机构研制成功一种效果很好的晚期癌症止痛药,我国制药人员也在研究这类药物,但未获成功。后来通过某种途径获得国外某一关键方法,终使自己获得成功。此一关键方法的知识及其载体(文献)就是情报。所以说情报是知识的一部分。被激活、被利用了的知识才属于情报。医学情报,则指获取所需的医学知识。在国外信息和情报没有区别,都用 Information 表示,我国过去用的“科技情报”,从 1992 年起国家科委通知都改用“科技信息”,以便与国际接轨。

4. 文献(Literature) 凡属于人类的知识,用文字、图形、符号、声频、视频等手段记录保存下来,可用以交流传播的一切物质形态的载体,统称为文献。概括地说,文献就是记录有知识的一切载体。它含有文献知识、文献载体、文献记录方式和文献符号系统四个方面。

文献知识:是文献的实质内容。不同学科、不同专业的知识,构成了不同内容性质的文献,如医学知识、农业知识。

文献载体:是文献的外在形态。知识被固定在不同的载体上,就形成了不同载体类型的文献,如纸张、光盘、磁盘、竹简、布帛、铜鼎等。

文献记录方式:是将知识固化在载体上的手段。知识以各种符号的形式,通过各种记录方式(写画、雕铸、印刷、拍摄、磁化、光刻等)依附在载体上形成文献。

文献符号系统:是将知识固化在载体上的表现形式,知识以文字、图形、代码、声频和视频信号等各种不同的符号及符号系统记录在载体上。

这四个方面相辅相成,缺一不可不能成为文献。

5. 文献与信息、知识、情报的关系

信息、知识、情报和文献之间的关系可以概括为：前三者的逻辑关系可以用图 1-1 三者关系图来表示。显而易见，知识和情报都属于信息这个大范畴，情报属于信息和知识这个小范畴。在一定条件下，三者可以相互转化。

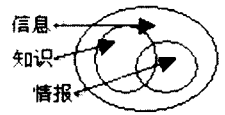


图 1-1 信息知识情报关系图

信息、知识、情报与文献之间既有区别又有联系。信息广泛存在于自然界和人类社会，其涵盖面最广，知识是经过人脑思维系统化了的那部分信息，而情报是经过激活而活化了的知识。文献是信息、知识、情报存储的方式，形成文献后才能被保存下来，才能进行传递，为后人所借鉴、利用，是物化了的信息、知识和情报。

二、文献类型

文献的类型，由于其外在形式和内容特征不同，可按下列标准进行划分。

（一）按对文献内容的加工深度分

1. 一次文献 即原始文献。它是指作者以本人的工作经验、科研成果为依据而撰写的并公开发表或公布的原始文献。不论撰写时是否参考或引用别人的资料，不论载体和出版类型，均为一次文献。专著、期刊论文、学位论文、研究报告、会议论文、专利说明书等都属于一次文献。一次文献是对知识的第一次加工。

2. 二次文献 通常称为检索工具，是人们把大量的、分散的、无序的一次文献收集起来，按照一定的方法进行加工、整理，使之系统化便于查找而形成的报道性、检索性的文献资料，是对知识的第二次加工。主要类型有目录、索引、文摘等。

3. 三次文献 是利用二次文献并在其指引下，对大量一次文献进行综合分析研究加以浓缩和提炼而写成的文献，是对知识的第三次加工。其类型有两种：一种是将大量发展成熟且长期有用的知识系统化，供查找和利用的参考工具书，如年鉴、百科全书、指南、手册等；另一种是就某些具体研究课题的当前情况和发展趋势而撰写的综述、述评、学科年度总结等。

从一次文献到二次文献的过程是一个由博而约，由分散到集中，由无组织到系统化的过程。从文献检索来说，一次文献是检索的对象，二次文献是检索的手段与工具。

（二）按文献载体形式分

1. 书写型文献 是指在印刷术尚未发明之前的古代文献和当今尚未正式付印的手写笔录，以及正式付印前的草稿。如古代的甲骨文、金石文、纸帛文、竹木文以及现今的会议录、手稿等。

2. 印刷型文献 又称纸质型文献，指以纸张为主要载体，以印刷为记录手段而产生的文献。包括石印、铅印和胶印等。

3. 缩微型文献 指以感光材料为载体，将纸张文献缩微复制在感光材料上而成的文献。包括缩微胶片和缩微胶卷。多用于存贮过期报纸、期刊、学位论文等。

4. 视听型文献 又称声像型文献，是指以磁性材料或感光材料为载体，采用录音、录像或摄影技术直接记录声音信息或图像信息而形成的文献。包括唱片、录音带、录像带、幻灯片、电影片等。

5. 数字化文献 指以数字化技术将文献储存在光、磁载体上，通过计算机或网络进行阅读的文献。可分为数据库文献和网络文献信息。

其中印刷型文献和数字化文献为当前科技文献的两大主流载体。

（三）按出版形式划分

1. 图书(Book) 是指一些记录的知识比较系统、成熟的文献。但编著和出版的时间较长，所论述的知识与期刊论文等文献相比有 3~5 年的时差。图书可分为供读者阅读的图书和供读者查阅的工具书，前者如单卷书、多卷书、丛书，后者如目录、索引、手册、辞典、年鉴、百科全书。按图书内容可分为文艺类、普及类、专业类图书。

在每一种正式出版的图书的版权页或其他明显部位标有一个国际标准书号(International

Standard Book Number,简称 ISBN),这是一种国际通行的出版物代码,代表某种特定图书的某一版本。ISBN 的唯一性和专指性使读者有可能利用 ISBN 号通过某些检索系统查询某种特定图书。

2. 期刊(Periodical, journal) 也叫杂志,是指一些记录的知识比较新颖、所含信息密度比较大的定期或不定期的连续出版物,一般都有固定的刊名、统一的出版形式。期刊按其出版周期(或刊期)有周刊、旬刊、半月刊、双周刊、月刊、双月刊、季刊、年刊之分;按其语种有中外文之分,外文又有西文、俄文、日文等之分;按其新旧和是否装订有现期期刊(简称现刊)和合订本期刊(简称过刊)之分;按纸张印刷质量和出版订购渠道又有原版和影印期刊之分。

国际上用于识别连续出版物(包括期刊)的编码是国际标准连续出版物编码(International Standard Serial Number,简称 ISSN)。每一种经过申请的出版物都可得到一个固定不变的标准号码。解决同名期刊辨识问题的另一种期刊编码系统是 CODEN 码(Code number)。

期刊通常刊登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,所以也称为学术期刊,是科研课题工作的主要文献源、信息源、情报源,约有 65% 的信息来源于期刊。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. 科技报告(Scientific & technical report) 是研究机构提交的围绕某一课题或项目所作的论证、总结、方案或介绍的报告。特点是内容详尽、专深,尤其具有参考价值。目前全世界每年出版的科技报告大约有 70 万件,主要有美国政府的四大套报告,即 AD(美国武装部队技术情报服务处)报告、AEC(美国原子能委员会)报告、NASA(美国国家宇航局)报告和 PB(美国政府出版局)报告。

4. 专利文献(Patent document) 指专利申请书、专利说明书、专利公报等与专利有关的一切文献。具有技术内容广泛、反映新技术快、内容翔实、标准化程度高等特点。

5. 会议文献(Meeting paper) 指在学术会议上宣读或交流的论文及会后汇集的论文集和专刊。其中国家级和国际学术会议论文质量高,专业性和科学性强,代表该学科技术领域研究的前沿。

6. 政府出版物(Government publication) 指政府部门及其所属的专门机构发表出版的行政性文件和科技文献。如法律法规、决议指示、技术政策等。

7. 学位论文(Dissertation) 指博士或硕士研究生在毕业前为获得学位而撰写的反映一定学术水平的论文。学位论文大多不公开出版发行,故不易搜集到,但历来被大学和专业图书馆所重视。

8. 标准文献(Standard document) 是由国家某一机构颁发的对工农业技术产品和工程建设的质量、规格、及其检验方法所作的各种技术规定的文件。标准具有一定的法律约束力,是从事生产、设计、施工的技术依据。

9. 技术档案(Technical archives) 是围绕某一工程对象而形成的有关技术文件、设计图纸、施工图纸、图表、照片、原始记录等。对医学科技来说,住院病历、科技实验数据记录等都是极有价值的档案材料,应重视保管和充分发挥作用。

10. 产品资料(Product material) 是对定型产品的性能、构造、原理、用途、使用方法和操作规程、产品规格等所做的具体说明。包括产品样本、产品标准、产品说明书、产品目录等。

11. 其他文献 如报纸、内部刊物、通报等,具有报道快、涉及面广、内容新等特点。

第二节 信息检索和文献检索

一、信息检索的概念

信息检索(Information retrieval)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

两个过程(Storage and Retrieval),是指将信息按一定的方式组织、存储起来,构成检索工具和根据用户的需要从检索工具中查找出所需信息两个方面。信息存储是指工作人员将大量无序的信息集中起来,根据信息源的外表特征和内容特征(如文献的标题、作者、来源和主题等),用特定的检索语言转化为一定的标识(如主题词、关键词、分类号和类目名称等),经过整理、分类、浓缩、标引等处理,使其系统化、有序化,并按一定的技术要求建成一个具有检索功能的工具或检索系统,供人们检索和利用。而信息检索是指运用编制好的检索工具或检索系统中的检索标识,查找出满足用户要求的特定信息的过程。信息检索的原理如下:

存贮过程→

信息检索(狭义): { 信息存贮:数据、事实、文献→分析→检索语言标引→排序→检索工具
检索工具→分析→检索语言→检索→数据、事实、文献

检索过程→

我们通常所讲的信息检索是指狭义的信息检索,即从检索工具或检索系统中,通过一定的检索途径或检索方式查找出需要的信息的过程。

为了保证文献信息能存得进、取得出,就必须使文献存贮所依据的规则与文献信息检索所依据的规则尽量做到一致,必须有统一的检索语言和名称规范作为存储人员和检索人员的共同依据。

二、信息检索的分类

信息检索(狭义)根据其检索对象的不同,可分为数据检索、事实检索和文献检索。

1. 数据检索(Data retrieval) 是指从贮存数据的集合中找出数值性数据的行为、方法或程序。数值性数据包括各种数字、图表、化学结构式、计算公式等。以特定的数据为检索对象,如临床实验室各种指标的正常值。

2. 事实检索(Fact retrieval) 是指从贮存的文献中查出关于某一事物发生与发展情况及相关信息的行为、方法或程序。以特定的事实为检索对象,如:2004年医学诺贝尔奖获得者是谁,何时何人在何种文献上首先报道乙型肝炎疫苗研制成功等。

3. 文献检索(Document retrieval) 是指向检索系统提交文献的外表特征或内容特征从某个文献集合(数据库)中查找出课题需要的文献信息的过程。

为了保证文献信息能存得进、取得出,就必须使文献存储所依据的规则与文献信息检索所依据的规则尽量做到一致,还必须有统一的检索语言和名称规范作为存储人员和检索人员的共同依据。

三、文献标引和检索语言

(一) 定义

1. 文献标引 就是对文献的外表特征和内容特征进行分析,并依据一定的标准或规则,用标识符号(检索语言)将其充分、准确、简明地表达出来的过程。为了使文献存贮得进去,检索得出来,文献标引和检索必须使用相同的语言标准。我们将这种统一的标引用语和检索用语的人工语言称为检索语言。

2. 检索语言 是应文献信息的加工、存储和检索的共同需要而人工编制的专门语言,是用来描述信息源特征和进行检索的一系列概念标识系统。检索语言可分为规范化语言和非规范化语言(自然语言)两类。

(二) 类型

目前,世界上的信息检索语言有几千种,依其划分方法的不同,其类型也不一样。下面叙述两种常用的检索语言划分方法及其类型。

按照标识的性质与原理划分:

1. 分类语言 是指以数字、字母或字母与数字结合为基本字符,采用字符直接连接并以圆点

(或其他符号)为分隔符,以基本类目为基本词汇,以类目等级的从属关系来表达复杂概念,以分类法(表)为分类体系和依据的一类检索语言。

以知识属性来描述和表达信息内容的信息处理方法称为分类法。著名的分类法有:《国际十进分类法》、《美国国会图书馆图书分类法》、《国际专利分类表》、《中国图书馆分类法》(简称《中图法》)等。其中《中图法》是我国使用最普遍的分类法。

分类语言能较好地体现学科的系统性,提示知识的派生、隶属与平等关系,便于检索者从学科或专业途径进行族性检索,并可按需扩大或缩小检索范围。

2. 主题语言 是指以自然语言的字符为字符,以名词术语为基本词汇,用一组名词术语作为检索标识,以词表为依据的一类检索语言。以主题语言来描述和表达信息内容的信息处理方法称为主题法。主题语言又可分为标题词、元词、叙词、关键词。

①标题词:是指从自然语言中选取并经过规范化处理,表示事物概念的词、词组或短语。标题词是主题语言系统中最早的一种类型,它通过主标题词和副标题词固定组配来构成检索标识,只能选用“定型”标题词进行标引和检索,反映文献主题概念必然受到限制,不适应时代发展的需要,目前已较少使用。

②元词:又称单元词,是指能够用以描述信息所论及主题的最小、最基本的词汇单位,经过规范化的能表达信息主题的元词集合构成元词语言。元词法是通过若干单元词的组配来表达复杂的主题概念的方法。元词语言多用于机械检索,如穿孔卡片等来标识信息。

③叙词:通常称主题词,是表达文献主题内容的经过规范化处理的名词术语。所谓人工规范化处理,即对文献中的同义词、近义词、多义词等加以规范,使得同一主题概念的文献相对集中在一个主题词下,同时在主题词表中采用参照系统间接反映主题概念之间与文献内容之间的关系,从而体现了主题词的单一性。例如:癌(非主题词)Cancer, tumor, carcinoma 见肿瘤(主题词)Neoplasms(详见第三篇第六章第一节)。采用主题词语言进行检索,能较好地实现较高的查全率和查准率,是提高检索技能的关键。适用于计算机和手工检索系统,是目前应用较广的一种检索语言。

④关键词:是指直接从文献标题、文摘及正文中抽出的有实质意义的,用来表达文献主题内容的关键性语词。关键词是一种非规范化(未经严格规范化处理)检索语言,属于自然语言范畴。比如癌(非主题词)Cancer, tumor, carcinoma。采用关键词检索的优点是能够及时反映文献的最新信息,用法简便,一般用户不经系统培训也容易掌握利用。因为用关键词不像主题词具有单一性,对同义词、近义词未经严格规范,也不设参照系统显示词与词之间的等级关系和相关关系,因此同一主题概念的文献标引相对分散,容易导致漏检和误检。查准率较好,查全率较低。

3. 代码语言 是指对事物的特征用某种代码系统来表示和排列事物概念,从而提供检索的一种检索语言。例如,根据化合物的分子式这种代码语言,可以构成分子式索引系统,允许用户从分子式出发,检索相应的化合物及其相关的文献信息。

按照表达文献的特征划分:

(1) 表达文献外部特征的检索语言

表达文献外部特征的检索语言主要是指文献的篇名(题目)、作者姓名、出版者、报告号、专利号等。将不同的文献按照篇名、作者名称的字序进行排列,或者按照报告号、专利号的数字进行排列,所形成的以篇名、作者及号码的检索途径来满足用户需求的检索语言。描述文献外表特征的检索语言可简要概述为:

{	题名——题名索引
	著者——著者索引
	文献编号 { 专利号索引 合同号索引 存取号索引
	其他——人名索引、引用文献目录等

(2) 表达文献内容特征的检索语言

表达文献内容特征的检索语言主要是指所论述的主题、观点、见解和结论等。

描述文献内容特征的检索语言可概述为：

{	体系分类语言——分类索引(目录)
	叙词语言(主题词)
	关键词语言
	其他语言——分子式索引、专利号索引等

在以上两类检索语言中,描述文献内容特征的语言与描述文献外表特征的语言相比较,在揭示文献特征与表达情报提问方面,具有更大的深度,标引和检索时,更需要标引人员和检索人员的智力判断。

第三节 书本式医学文献检索工具

一、国内出版的主要医学文献检索工具书

1. 《全国报刊索引·自然科学技术版》是综合性题录式的索引刊物,月刊,由上海图书馆编辑出版。它汇编中央和各省、自治区、直辖市出版的报纸和期刊所登载的文章目录,具有收录范围广、报道量大和时差短等特点。它除了收录医学报刊外,还收录了与医学相关的边缘学科报刊,其收录量和涉及的学科范围均多于其他中文医学检索工具。

该刊的编排结构为:编辑说明、分类目录、正文题录、作者索引和引用报刊一览表。其著录内容包括:题录号、文献题名、著者姓名及所在单位名称(最多著录前三名,其余作者用“…”表示)、刊名、出版年、卷、期和起止页码。

2. 《中文科技资料目录·医药卫生》简称《中目医》,是大型题录式科技检索刊物《中文科技资料目录》的分册,月刊,由中国医学科学院医学信息研究所编辑出版。主要收集国内医学和与医学有关的期刊、汇编、学术会议资料等,年报道量约5万条。是查找国内医学文献的主要检索工具,时差约一年。

该目录采用以学科分类为主、主题索引为辅的编排方法,由编辑说明、分类目次、正文题录、本期学科分类类名索引、主题索引和附录组成。其题录著录格式包括题录顺序号、文献题名、著者姓名及所在单位名称(只著录第一著者,其余用“…”表示)、期刊名称、期刊编辑或主办单位、出版年(期)和起止页码。

3. 《国外科技资料目录·医药卫生》简称《外目医》,是国外科技资料的大型题录式检索刊物,医药卫生为其分册之一,月刊,由中国医学科学院医学信息研究所编辑出版。收编英、法、德、俄、日五种语言的500余种期刊。著录格式包括题录顺序号、中译题名、著者姓名、期刊或其他资料名称、年、卷(期)、起止页、文种和摘要、译题单位及索取号。

4. 《中国医学文摘》是由国内有关医疗单位分别编辑出版的医学文献检索体系,共18个分册。它用文摘、简介、题录的形式全面报道国内公开发行的医药卫生期刊。以论著、病例报告、科研成果、先进技术等文献作为摘录重点。著录格式包括文摘顺序号、文献题名、著者、期刊名称、期刊

年、卷(期)、起止页码、文摘、附表和参考文献数量、文摘员。

5. 《中国药学文摘》月刊,由中药资料电脑检索中心、国家医药管理局医药技术情报所和中国医学科技情报研究所编辑。本刊刊载我国公开发行的药学期刊、医药杂志、医学杂志、医药院校学报,以及化工、植物、微生物杂志近 200 余种。以文摘为主,也以提要、简介和题录等形式报道。著录格式包括中国图书分类号、药学期刊号、文摘号、文献题目、作者、作者所在单位、期刊名称、年卷期及起止页码、文摘、专利文摘与专利号、文摘员。

6. 《国外医学》由中国医学科学院情报研究所组织全国 35 所医学院校、科研单位、医学情报机构编辑的大型医学情报系列刊物,综合报道国外医学领域的研究进展。按专业划分为 47 个分册。每期均刊有综述论文、译文和文摘。综述论文多为国内专家撰写,文后附有大量参考文献,也有的是全文翻译国外较好的综述论文,文末附有原文出处。不仅在专业方面提供国外动态综述性的情况,而且提供国内外如何写好综述的范例。文摘部分可提供重要的国外医学研究信息。

二、国外出版的主要医学文献检索工具书

1. 美国《医学索引》(Index Medicus,简称 IM) 是当今世界上最著名的大型生物医学文献题录型检索刊物,1879 年创刊,由美国国立医学图书馆(National Library of Medicine,简称 NLM)编辑出版。IM 具有以下特点:收录范围广,收录全世界 72 个国家、41 种语言的生物医学相关期刊 3400 余种,年收录量约 40 万篇,英文文献占 75% 以上。时差短,一般时差 2~3 个月,最短时差仅 1 个月。1964 年起,NLM 建立了医学文献分析和检索系统 MEDLARS(Medical Literature Analysis and Retrieval System),采用计算机技术加工文献,从而加快了出版报道速度。内容全面,但对临床稍有侧重。

IM 为月刊,每年出版 12 期为一卷。每期分两部分:Part I——Subject Section 主题部分,Part II——Author Section 著者索引,Bibliography of Medical Reviews 医学综述题录。当一年月刊出齐后,NLM 另外单独出版该年的累积本,称为累积医学索引(Cumulated Index Medicus,简称 CIM)。CIM 结构编排与月刊 IM 完全相同,回溯检索时应尽量利用已出版的 CIM,以节约检索时间,提高检索效率,它完全可以代替当年的月刊本。指导 IM 主题检索的《医学主题词表》(Medical Subject Headings),每年随 IM 的第一期和 CIM 出版。随第一期出版的还有《收录期刊一览表》(List of Journals Indexed in Index Medicus)。

2. 荷兰《医学文摘》(Excerpta Medica,简称 EM) 是一种文摘型医学英文检索工具,1947 年创刊,由总部设在荷兰阿姆斯特丹的医学文摘基金会(The Excerpta Medica Foundation)编辑出版,故称荷兰《医学文摘》。EM 涉及面广,文摘量大,收录 110 多个国家和地区的 4000 多种医学刊物。内容包括临床、药物、基础生物医学研究、卫生等。EM 按医学专门研究领域划为各分册出版,创刊时仅 8 个分册,至今已有 42 个分册。各分册每年出版 1~4 卷,每卷可有 10、8、7、6 期不等。为了更好地方便读者使用 EM,1984 年出版了《医学文摘分类与索引系统指南》(A Guide to the Excerpta Medica Classification and Indexing System),其中刊载各分册的全部分类目次表以及“概念字顺表”,包括约 5000 个生物医学词汇,每个词汇后列出相应的分册号,可帮助检索者了解某个主题应选哪种检索途径、哪一分册最适宜。EM 已建立了 EMBASE 数据库,且已进入 DIALOG, DATASTAR, DIMDI, LEXIS/NEXIS, Ovid Online 和 STN 等国际联机检索系统。在 DIALOG 系统中 EM 为 72、172、173 三个文档,可检索自 1974 年以来的该文摘所载内容。1996 年出版了光盘 19 种。目前 EM 也进入了 Internet 网,从剑桥文摘所(Cambridge Scientific Abstract)的 Internet 数据库资源中已能获取 EM 的 15 个分册的文献(网址:<http://www.excerptamedica.com>)。

3. 美国《生物学文摘》(Biological Abstracts,简称 BA) 创刊于 1926 年,由《细菌学文摘》和《植物学文摘》合并而成,现由美国生物科学信息服务社(Bioscience Information Service,简称

BIOSIS)编辑出版。BA是生命科学的重要文摘型检索工具,收录110多个国家和地区的7000多种期刊,尤其是欧美各国有关生物学、医学、农业方面的文献。在医学方面,报道的重点为神经系统、肿瘤及肿瘤因子、分子生物学、药理学、心血管系统、遗传、公共卫生、生态学、免疫学、微生物、毒理等方面,尤其在生物实验研究方面有独到价值。BA目前为半月刊,半年为一卷(12期),每年出24期(2卷),从1998年第105卷起有很大幅度的变化,每期分为文摘与索引两部分。

4. 美国《化学文摘》(Chemical Abstracts,简称CA)创刊于1907年,由美国化学会的化学文摘社(Chemical Abstracts Service,简称CAS)编辑出版。创刊初期为半月刊,一年一卷;1962年改为双周刊,半年一卷(13期),一年两卷;1967年66卷起改为周刊,每年两卷共52期,并沿用至今。CA每周出版一期,每年52期,分为两卷,每卷26期,称为期索引,是CA的核心。每卷出齐后随即出版一套卷索引。每隔10年或5年出版一次累积索引。目前已出版了13次累积索引。CA的载体形式除了印刷型外,还有缩微胶片、机读磁带、光盘及联机版式。现在通过Internet可在许多网站查询CAS提供的数据库。

为了帮助用户更好地使用CA,CAS还出版有索引指南(Index Guide,简称IG)、资料来源索引(CAS Source Index)和化学物质登记号手册等指导性索引。在报道的所有文献中超过1/3的文献为生物医学文献。CA所收录的1000种核心期刊中,有关生理、生化、微生物、环境医学、药理、临床医学等学科的期刊占1/3,包含了最重要的医学核心期刊,因此CA已成为查找医学文献最重要的补充检索工具(网址:<http://www.cas.org>)。

5. 美国《科学引文索引》(Science Citation Index,简称SCI)1961年创刊,由美国科学情报所编辑出版,为一种大型多学科综合性检索工具,也是一种比较科学的检索体系。它收录了世界上约43个国家和地区3000多种重要科技期刊,年收集论文数量为60万条以上。内容涉及化学、医学、生物学、农业、地球科学等近百个学科领域,其中生命科学及医学、化学、物理学所占比例最大。由于编制方法独特,通过揭示引证与被引证的关系,可以沟通当年、几年、几十年甚至上一世纪与当前作者和学科之间的关系,已越来越被人们所认识与利用。目前,为了制定科学指标,调查分析国力,一些国家如美国、俄罗斯、德国、英国、匈牙利等都利用此种工具进行论文统计与分析。

第四节 数据库的结构和类型

数据库(Database)是因特网上重要信息资源类型。数据库由若干文档(Files)组成,是文献的机读记录的有序集合。数据库存储于磁盘、磁带或光盘等载体上,网络数据库则是存储在网络检索系统和网站服务器的硬盘中,并能在网络环境下通过计算机进行检索。一个数据库相当于一种书本式手工检索工具。

一、数据库的结构

文档(Files)是数据库的组成部分。按其编排结构与功能分为顺排文档和倒排文档。

(一) 顺排文档(Sequential File)

这是数据库中的主文档,是大量记录(Records)的集合,是按每条记录输入顺序编号(称“存取号”)和排列的目录式文档。检索结果的信息都来自顺排文档。它相当于印刷型手工检索工具的正文题录或文摘部分。

1. 记录(Records)是数据库中一个完整的信息单元,由一系列字段(Fields)构成。一条记录代表一篇文献,如一篇期刊论文、文献综述、会议论文、专利说明书等。

2. 字段(Fields)是著录文献各项的名称及其内容,是描述文献内外特征的各项标识内容。各字段都有相应的字段标识符,如题名(TI)、著者(AU)、地址(AD)、出处(SO)、主题词(MeSH)、主要主题词(MMH,MJME)、关键词(KW)、特征词(TG)、出版年(PY)、语种(LA)、摘要