

高等医学院校卫生事业管理专业教材

医学文献检索

(第2版)

主 编 谢志耘



北京大学医学出版社

高等医学院校卫生事业管理专业教材

医学文献检索

(第2版)

主 编 谢志耘

编 委 (以姓氏拼音排序)

高 琴 李春英 刘春艳

沈 霞 张燕蕾

北京大学医学出版社

YIXUE WENXIAN JIANSUO

图书在版编目 (CIP) 数据

医学文献检索/谢志耘主编. —2 版. —北京:
北京大学医学出版社, 2010. 1
ISBN 978-7-81116-873-0

I. ①医… II. ①谢… III. ①医学—情报检索—高等
学校—教材 IV. ①G252.7

中国版本图书馆 CIP 数据核字 (2009) 第 231293 号

医学文献检索 (第 2 版)

主 编: 谢志耘

出版发行: 北京大学医学出版社 (电话: 010-82802230)

地 址: (100191) 北京市海淀区学院路 38 号; 北京大学医学部院内

网 址: <http://www.pumpress.com.cn>

E - mail: booksale@bjmu.edu.cn

印 刷: 北京瑞达方舟印务有限公司

经 销: 新华书店

责任编辑: 韩忠刚 责任校对: 杜 悦 责任印制: 郭桂兰

开 本: 787mm×1092mm 1/16 印张: 14.25 字数: 373 千字

版 次: 2010 年 1 月第 2 版 2010 年 1 月第 1 次印刷 印数: 1-8000 册

书 号: ISBN 978-7-81116-873-0

定 价: 19.90 元

版权所有, 违者必究

(凡属质量问题请与本社发行部联系退换)

前 言

自 20 世纪 90 年代以来,随着计算机和网络通信技术的迅猛发展,特别是因特网技术在社会经济生活中的广泛应用,人类已经不知不觉地迈入了信息化时代。这个时代使人们深切地感受到:知识就是力量、信息就是财富。信息资源已与物质资源、能源资源一起被全世界公认为人类社会的三大支柱之一。全球知识经济的快速发展正是一个很好的例证。在这个发展过程中,信息资源发挥了重要的作用,一方面信息成为知识创新的因素和原动力,另一方面,为了有效获取和利用信息,必须促进信息科技的发展,这又进一步推动了知识经济的加速发展。因此,把握了信息和信息资源,也就是把握了现在和未来。

1995 年,国家提出实施“科教兴国”的战略,要进行教育创新和改革,核心目标就是要不断造就大批具有丰富创新能力的高素质人才。一个具有创新能力的跨世纪人才必须具备多种素质,其中信息素养是不可或缺的重要组成部分。所谓信息素养是指使用计算机和信息技术高效获取、正确评价和善于利用信息的能力。在科技高速发展的今天,只有善于获取和利用信息的人,才有可能在前人的基础上进行创新活动。另一方面,终身学习、能力导向学习和开放学习日渐成为新的教育理念的重要内涵。为满足知识创新和终身学习的需要,发达国家纷纷将信息素养教育,作为培养 21 世纪人才的重要内容。医学文献检索课程是信息素养教育的重要组成部分,是培养学生的信息意识、掌握用手工方式和计算机方式从浩瀚的文献信息海洋中获取知识和情报的一门科学方法课,是提高学生自学能力和独立研究问题能力的工具课。

在高等学校开设文献检索课程已经有 20 余年的历史,其间课程的讲授内容、授课方式都发生了巨大的变化,这都源于信息技术的发展。首先,作为信息资源集散地的图书馆从传统的以纸本收藏为主逐步过渡到以电子馆藏和网络服务为主;其次,查找文献信息的工具和方法也发生了巨大的变化,从传统手工检索过渡到计算机和网络检索。这就使得文献检索这门课程始终处于动态变化之中。

本教材是在 2002 年出版的《医学文献检索》第 1 版的基础上修订而来。由于文献资源、文献检索工具及其使用方法发生了巨大的变化,因而本版教材相比第 1 版教材在内容上已有 95% 以上的内容更新。

本教材是专为高等医学院校卫生事业管理专业的学生而编写。在第 1 版的基础上,对教材内容、结构进行了重新规划和设计,强调理论联系实际,具有深入浅出、文字简明等特点。本教材内容仍划分为 9 个章节,包括医学文献检索基础、图书馆信息资源利用、电子图书、电子期刊以及文摘数据库检索、引文数据库检索、特种文献检索、因特网医学信息资源检索以及医学文献的综合利用。本教材立足于在职人员继续教育的需要,教材编写力求结构合理、内容丰富、实用性强。根据本课程实践性强的特点,书中采用了较多的实例,并以图文并茂的方式表达,便于读者理解与掌握,也便于学生自学。

由于编者学识和水平所限,加之时间仓促,书中可能有不少错讹或不妥之处,祈望各位同行和专家批评指正。

谢志耘

2009 年 12 月

目 录

第一章 医学文献检索基础..... 1	第三节 图书馆信息服务 45
第一节 文献与文献检索..... 1	一、电子阅览服务 45
一、文献基础知识..... 1	二、定题服务 47
二、医学文献的特点..... 4	三、情报调研 47
三、文献检索基础知识..... 5	四、科技查新服务 48
四、文献检索系统..... 7	五、学科化服务 49
五、文献检索的意义 10	六、个性化导读服务 50
第二节 文献信息资源及其类型 11	测试题 50
一、文献信息的类型 11	第三章 电子图书检索 54
二、电子文献资源概述 14	第一节 电子图书概述 54
三、医学电子文献资源类型及 产品 16	一、电子图书的概念 54
第三节 文献检索语言 19	二、电子图书的发展 55
一、检索语言概述 19	三、电子图书的类型 55
二、检索语言的类型 19	四、电子图书的特点 56
三、医学文献检索语言 23	第二节 中文电子图书 58
第四节 文献检索方法、途径与 步骤 25	一、读秀电子图书 58
一、文献检索的基本方法、途径与 步骤 25	二、书生之家电子图书 64
二、计算机文献检索方法与步骤 ... 27	三、其他中文电子图书 66
测试题 31	第三节 外文电子图书 67
第二章 图书馆信息资源利用 35	一、NetLibrary 电子图书 68
第一节 联机公共书目查询系统 (OPAC) 35	二、OVID 电子图书 (Book@Ovid) 70
一、OPAC 简介 35	三、其他外文电子图书 73
二、OPAC 主要功能 36	测试题 74
三、OPAC 的一般使用步骤 37	第四章 文摘数据库检索 76
四、OPAC 检索实例 38	第一节 中国生物医学文献服务系统 (SinoMed) 76
第二节 图书馆传统服务 40	一、系统概况 76
一、书刊借阅服务 40	二、系统功能特点 77
二、馆际互借服务 42	三、检索技巧和方法 78
三、文献传递服务 43	四、检索结果的处理 83
四、文献复制服务 44	五、检索举例 84
五、代查代检服务 44	第二节 PubMed 数据库 85

三、检索方法与技巧	87	四、Wiley InterScience 电子期刊	
四、检索结果的处理	90	全文数据库	131
五、检索举例	91	测试题	133
第三节 EMBASE.com 数据库	92	第六章 引文数据库检索	135
一、概况	92	第一节 引文索引概述	135
二、检索运算符及运算规则	93	一、基本概念	135
三、检索方法与技巧	93	二、引文索引编制原理及方法	136
四、其他功能	96	三、引文索引的功能和局限性	137
五、检索结果的显示与输出	96	第二节 中国科学引文数据库	137
六、检索举例	97	一、数据库概况	137
第四节 CINAHL 数据库	98	二、检索方法与技巧	138
一、数据库概况	98	三、检索结果的处理	140
二、数据库检索规则	98	四、检索举例	141
三、检索技巧和方法	98	第三节 Web of Science 数据库	142
四、检索结果的显示与输出	101	一、数据库概况	142
五、检索举例	103	二、检索运算符及运算规则	143
测试题	103	三、检索方法与技巧	144
第五章 电子期刊全文数据库检索	106	四、检索举例	151
第一节 中国期刊全文数据库	106	测试题	151
一、数据库概况	106	第七章 特种文献检索	153
二、检索方法与技巧	107	第一节 专利文献检索	153
三、检索结果的处理	111	一、专利的基本知识	153
四、检索举例	112	二、专利文献的基本知识	157
第二节 中华医学会期刊全文数据库		三、国际专利分类系统	159
.....	113	四、中国专利文献检索	161
一、数据库概况	113	五、德温特专利检索	162
二、检索方法与技巧	113	第二节 学位论文检索	164
三、检索结果的处理	117	一、学位论文概述	164
四、检索举例	117	二、中国学位论文数据库	165
第三节 Elsevier 电子期刊	118	三、PQDD 博士学位论文数据库	
一、数据库概况	118	167
二、检索方法与技巧	119	第三节 会议文献检索	169
三、检索结果的处理与个性化服务		一、会议文献的基本知识	169
.....	121	二、中国医学学术会议论文数据库	
第四节 其他电子期刊	124	170
一、中文科技期刊数据库 (全文版)		三、ISI Proceedings 数据库	171
.....	124	测试题	172
二、Ovid 电子期刊全文数据库	125	第八章 因特网医学信息资源检索	175
三、ProQuest 电子期刊全文数据库		第一节 网络信息资源概述	175
.....	128	一、网络信息资源的定义	175

二、网络生物医学信息资源的特点.....	176	第九章 医学文献的综合利用.....	200
三、网络生物医学信息资源的类型.....	177	第一节 课题资料的搜集、整理与分析.....	200
四、网络生物医学信息资源的分布.....	179	一、课题资料的搜集.....	200
五、网络生物医学信息资源的评价.....	180	二、文献信息的甄别、评价与整理.....	202
第二节 搜索引擎.....	180	三、课题资料的分析.....	204
一、搜索引擎简介.....	180	第二节 参考文献管理软件.....	205
二、综合性搜索引擎.....	182	一、NoteExpress 概述.....	205
三、医学专业搜索引擎.....	187	二、NoteExpress 的安装与使用.....	205
第三节 开放获取资源.....	189	三、数据库文献的导入.....	208
一、开放获取的发展历史.....	190	四、辅助论文写作.....	209
二、开放获取的定义.....	190	第三节 医学论文写作.....	210
三、开放获取的形式.....	190	一、概述.....	210
四、网络开放获取学术资源.....	191	二、医学论文的类型.....	211
第四节 常用医学网站.....	194	三、医学论文写作步骤与方法.....	212
一、国外生物医学网站.....	194	四、医学论文撰写格式和要求.....	213
二、国内生物医学网站.....	196	测试题.....	217
测试题.....	197	参考文献.....	219

第一章 医学文献检索基础

学习目标

1. 掌握文献的定义，了解文献、信息、知识、情报之间的关系。
2. 熟悉文献的构成。
3. 了解医学文献的特点。
4. 掌握文献检索的定义，了解文献检索的类型和原理。
5. 熟悉文献检索系统的定义、组成要素、内容结构、类型和评价因素。
6. 掌握文献信息类型和医学电子文献资源类型。
7. 熟悉电子文献资源概念，了解电子文献资源特点。
8. 掌握检索语言的定义和类型。
9. 熟悉主要的医学文献检索语言。
10. 熟悉文献检索的基本方法、途径和步骤。
11. 掌握计算机文献检索方法与步骤。

第一节 文献与文献检索

一、文献基础知识

(一) 信息、知识与情报

文献与信息、知识、情报密不可分。在了解文献的概念之前，先了解一下信息、知识和情报的涵义。

究竟什么是信息？众说纷纭，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。我们认为，信息的概念可以这样来定义：信息是客观世界中各种事物的变化和特征的最新反映以及经过传递后的再现。信息是通过一定的物质载体形式反映出来的，是事物存在的状态、运动形式、运动规律及其相互关系、相互作用的表象。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。

关于知识的定义,汉语词典中解释为: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。例如根据某一症状、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。

情报最初产生于军事领域,主要是将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的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。

(二) 文献的概念

“文献”一词,在我国有几千年的历史,最早源于春秋,意指典籍(文)和贤者(献)。其后,“文献”之义几经变化。比较有权威的定義主要包括:《文献情报术语国际标准(草案)》(ISO/DIS5127)的定义,即“为了把人类的知识传播开来和继承下去,人们用文字、图形、符号、声频、视频等手段将其记录下来,或写在纸上,或晒在蓝图上,或摄制在感光片上,或存储在磁盘上。这种附着在各种载体上的记录统称为文献”;和《中华人民共和国国家标准·文献著录总则》对文献的定义,即“文献是记录有知识的一切载体”。这两个定义基本上揭示了文献的特征,但“各种载体”和“一切载体”比较笼统,有的载体是固态的,有的载体是动态的(如借助声波传递信息)。因此载体又被分为存储型载体和传播型载体。文献应属于存储型的固态载体,如印刷件、缩微制品、磁盘、光盘、硬盘、优盘等,而不应包括可承载和传递同样信息的电话、语音信箱、图文电报、电子公告板、网络等瞬时信息的附载物。

综上所述,文献可被认为是:“凡人类的知识或信息,以文字、图形、代码、符号、声频、视频等形式记载到不同固态物质载体上面形成的一切记录,都统称为文献”。文献是本课程的主要研究对象。

(三) 信息、知识、情报、文献之间的关系

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。

而文献的定义可以看出,文献则是被物化了的知识记录,是被人们所认知并可进行长期管理的信息。情报是人们为解决特定问题而被活化了的更为高级、更为实用的知识。情报蕴含在文献之中,但不是所有文献都是情报,而所有情报几乎都是知识。文献又是贮存传递知识、情报和信息的介质。知识是文献的实质内容,载体是文献的外在形式,而记录是联系知

识与载体的手段。

长期以来,人们习惯于从文献中获取信息、知识和情报,文献已成为人们获取信息、知识和情报的主要方式。因此,文献是一种重要的信息源。但是,文献并不是唯一的信息源。除了文献信息外,还有非文献信息源。非文献信息源又可分为实物信息源和口头信息源两种。前者包括实物、样品、展览等,后者包括交谈、会议、广播等。

文献是人类认识自然和社会过程的知识记录,是人类物质文明不断发展的产物,印下的一串串足迹可以再现历史、鉴往昭来;它凝结着数千年来人类点滴积聚的智慧和成果,所蕴含的知识、信息、情报,是人们读书治学、开展科学研究活动和从事其他工作的智慧源泉,可以使知识增值;是启迪后人开创未来的丰富宝藏;它可以满足人们的文化需求,使人们从中获取有益的启迪与精神享受,是精神财富的重要组成部分。

(四) 文献的构成

现代文献由四要素构成:文献信息、文献载体、符号系统和记录方式。文献信息是文献的内容,符号系统是信息的携带者,载体是符号赖以依附的“寄主”,而记录方式则是代表文献的符号进入载体的方法和过程,四要素缺乏任何一种都不可能形成文献。

1. 文献信息 这一要素是指文献载体上所承载的信息、知识内容。信息与知识是构成文献的主体,没有信息、知识内容的任何形式的载体,只能称是物质,而不能称为文献。任何形式、类型的文献都必须首先以具有一定的信息、知识内容为前提。

文献信息属于信息,除具有信息的一切性质和功能外,文献信息也有自身的特点。表现为:①文献信息是经人的一系列加工(筛选、归纳、整理)后记录下来的信息(Recorded Information),不是指文献符号系统本身的信息,也不是指文献载体本身的信息;②文献所表达的信息内容虽然与符号本身没有必然联系,但文献信息的传递是通过人工符号系统——文字、标识符、声像信号来实现的,因而对文献信息的摄取方式和吸收的程度必然受到这种人工符号的制约;③文献信息是一种相对固化的信息,如纸质印刷品所传递的信息是无法变动的;可擦写光盘和磁盘虽然可以用修改、删除、增入等手段更新数据,但一旦变动后又处于静态之中。文献信息的固态化是文献易老化的原因,它不能随外界的变化而变化。为了克服由此造成的弊端,需要不断更新文献;④文献所传递的信息是人对客观世界的反映,因而不一定完全符合客观世界表现出的信息内容,这种“歪曲”、“畸变”、“失真”的程度因人们的认识水平、立场观点、方法和时代因素的不同而异;⑤文献信息既然是人工记录下来的附载于物质实体上的信息,那么这种物质实体就是文献信息的“外壳”。两者的关系是既具有不可分割性(没有文字的纸不成其为书),又具有相对独立性,即信息内容不会因载体形式的改变而改变,或者说不同的信息载体可以传播同一内容的信息,这一特点告诉我们信息资源的开发涉及载体资源的开发。为了获取或传播同一信息,人们应该选择信息载体和传播方式,以便更方便、更有效、更经济地开发和利用文献信息。

2. 文献载体 信息内容必须依附于载体而存在。与其他信息载体不同的是,文献载体必须适应于文献符号和相应的记录方式,同时又要有利于传播、整理和长期的保存,不符合前一要求的就不称其为文献,不符合后一要求,文献就不能随时代的发展而发挥出应有作用,就会被淘汰、消失。因此文献载体的材质在人类文明的演进过程中一直在“新陈代谢”。从古至今,文献载体经历了从非人工材质的泥版、岩石、石板、兽骨、木板、竹片、兽皮、树叶、桦树皮等到各种人工材质,如无机材质(陶、砖瓦、瓷、玻璃等)、金属材质(青铜、铁、铝、金、银等)、高分子材质(帛、纸、胶片、醋酸纤维等)、复合材质(磁带、光盘

等)的转变。其中纸质文献具有价格低廉、质地柔软、易于书写、携带和收藏等其他一些载体所无法比拟的性能而成为文献家族的主干。印刷术的发展,装订技术的进步,以及社会对文献信息需求的增加,使印本文献产量和类型出现惊人的发展。期刊、图书、报告、会议录、政府文献、学位论文、标准、产品样本、专利、档案构成印本文献的主体和精华。而在科学技术高速发展的今天,因为其他介质的文献陆续应运而生,如缩微文献、音像文献、磁盘和光盘文献等,纸质文献又开始变得不便于信息的快速传输、高效查阅和高密度存贮了。可见,从文献产生之日起,它便开始以适于自己生存的载体形态来寻求自己的立足之地。各种形态的文献在接受自然与人类社会长期的实践与应用过程中接受淘汰和选择。

3. 文献的符号系统 文献中的符号系统指图画、文字、公式、图表、编码、声像和电磁信息等。图画是最早的文献信息符号,在文字出现以前,人类就用图画来表达精神信息,图画的独到之处在于较强的直观性,读画人(信宿)无需接受文字训练就可与作画人(信源)沟通。此外,图画让人们保持注意力和记忆力的作用比文字更大,因此以图代文,让人们在最短时间内有效地获取最大容量的信息成为今天一切宣传媒介设计所注重的方式。文字,是有声语言信息的书面表达形式,是由图画演变而来的。图画对于表达抽象思维信息的无能为力可由文字来弥补。文字用一套书写符号来记录、表达语言要素(词、词素、音节、音素等)。但文字的特点——高度概括性、使用的随意性、音义的双重性也往往导致只可意会、无法言传的模糊语言,导致词不达意、一音多义、一语多音的现象,给信息交流造成误解。这说明人工符号永远只能接近表述信息的原意。声像信息指留在唱片、录音带上的声频信息和摄在胶卷、录像带上的视频信号。电磁信息符号是指计算机可读磁盘或光盘上的信息符号,由二进制的0和1构成,这些信息符号不是我们感官直接提取的,必须通过计算机解读。

4. 文献的记录方式 文献记录方式是指将表达信息的符号系统通过特定的人工记录方式使其附着于一定的文献载体上。记录方式种类很多、职能各异,按记录手段分有写画、雕刻、印刷、摄制、录音;按记录方法分可有手工记录、机械记录、光记录、电记录、声记录和磁记录。其中最常见的是印刷。

随着科技的发展,文字记录可以转变为数据记录,并用电子方法存贮到磁介质上去。光学字符识别(Optical Character Recognition, OCR)技术就代表了这种崭新的技术,它是指通过光学方法对字符、标记表示的书面数据进行自动识别,转换成机器可以处理的信息的过程。通过OCR可以实现纸质文献信息向电子信息的转换。

二、医学文献的特点

医学文献是对医学研究成果的有效记录,也是对医学知识的公开展示。了解医学文献的特征,有利于医学工作者及相关人员更好地利用医学文献。

1. 数量庞大 医学文献是整个科技文献中的重要组成部分,从整个世界科技出版物来看,医学文献约占整个科技文献的四分之一。据联合国教科文组织统计,全世界每年出版的图书80万种以上,平均每15年翻一番,科技期刊在10万种以上;每年产生约600万篇文献,100万件专利,其中生物医学文献占30%~40%。可见医学文献数量非常庞大,占有学科之首。

2. 载体多样化 目前,医学文献的载体除了传统的印刷型外,还有视听型(如录音带、录像带、电影、幻灯等)、缩微型(如缩微胶片、缩微胶卷、缩微卡等)、机读型(如磁盘、

磁鼓、光盘、优盘等)等。而且磁介质文献载体和光盘已广泛用于记录文献。随着科学技术的日新月异,新的文献载体亦还在不断出现。有人认为将来印刷型文献可能会被其他类型的文献所取代。不过,从目前来看,由于印刷型文献与其他类型的文献各有自己的长处和短处,加之其他因素,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. 多语种化 各种资料表明,全世界出版的文献语种正在不断增加。在过去,绝大多数医学论文仅用少数大语种发表。随着科学的发展,特别是发展中国家的经济文化进步,医学文献的语种急剧增多。据统计,现在只有60%左右的医学文献是用英、德、俄三种文字发表,余下40%的文献则以60余种语言发表。文种的增加丰富了医学文献,拓宽了研究空间。

4. 学科交叉、出版分散 现代科学技术综合交叉、彼此渗透,以至同一篇文献可以用不同形式、不同文字、在不同范围内多次发表。如会议论文、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他形式出版。一般专题范围的文章都刊登在本专业期刊上,但也可以刊登在边缘期刊或综合性刊物上。据统计,在直接相关的专业杂志上发表的文献只占50%,另外一些则发表在其他间接相关的专业杂志上。各学科的交叉渗透,致使大量的重要医学文献分散在相关乃至不相关的期刊上,仅凭一、两种医学期刊已无法掌握本专业的全部文献,只有靠编制检索工具将同类文献的线索集中在一起以便查阅。

5. 知识信息更新加快 科学技术的迅速发展,促使知识信息的陈旧速率加快,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。在研究活跃的生命科学和医学领域,医学文献的代谢更是频繁,大学生毕业后10年在学校里学的知识80%以上是陈旧过时的,所以当今的每位医学工作者必须以终生教育取代一次性教育。

6. 交流传播速度加快 现代交通、通讯、光学和印刷技术的发展,特别是现代通讯技术和电子计算机相结合应用技术的飞速发展,以及文献信息载体的电子化等在信息交流传播中的应用,为文献信息的快速传递与交流提供了非常便利的条件。文献信息的用户,可以通过信息交互网在瞬间即可获得所需要的文献情报信息。如对“甲型H1N1”流感疾病的预防控制或相关疫苗研究成果的情报信息在报刊上一公布,就会立刻传播到世界各地。

7. 电子化发展趋势 虽然目前各种载体形式的医学文献同时并存,但可以预见,随着Internet网络技术的普及推广使用,将从根本上改变信息存取与传播的方式,电子型文献也会越来越受到人们的关注。数字化图书馆(digital library)、无纸社会(paperless society)等新概念应运而生,计算机、网络将成为传播文化知识、获取信息资源的重要手段。电子期刊、电子书籍、电子词典、电子图书馆将成为医学知识信息的重要来源。

三、文献检索基础知识

(一) 文献检索的概念

文献检索(Literature Retrieval)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”。

(二) 文献检索的类型

1. 按照检出结果的形式划分 按照检出结果的形式不同,文献检索一般可划分为下面几种类型:

(1) 书目检索:指检索出来的结果是关于文献的题录信息,即获得的是关于文献的题名、作者、摘要、出处等此类的文献线索信息,检索者可根据这些文献线索进一步获取文献原文。如:美国《医学索引》、荷兰《医学文摘》等都属于书目检索工具。

(2) 全文检索:是指检索出的结果直接为文献的全文内容,或者是按照用户需求,检索出有关的句子、章节或段落。全文检索是在计算机检索技术的支持上发展起来的。如:《中国期刊全文数据库》、《ProQuest 保健、医学与药学全文数据库》等,都是全文检索数据库。

(3) 引文检索:指检出结果为文献被引用的情况,主要用以分析和评价文献质量。如《科学引文索引》(Science Citation Index, SCI),可检索出某著者或文献被其他著者或文献引用的情况;又如《期刊引文分析数据库》(Journal Citation Reports, JCR)是一种关于期刊评估的基本的、全面的资源工具,可查询期刊的影响因子(IF),可以根据期刊引用数据反映期刊的重要程度。

2. 从情报检索角度划分 有观点认为广义的文献检索与情报检索(information search)的意义相同,认为文献检索就是情报检索。根据检索对象的不同,情报检索又分为文献检索、数据检索和事实检索。

(1) 文献检索:这里的文献检索是指以文献为检索对象,其检索结果是文献线索或具体的文献。从性质上说,这种文献检索属于一种相关性检索,即它只需检出与用户所需相关的文献供用户参考。例如:检索“有关生态关系社会的发展与人类健康方面”的文献,就是属于文献检索范畴的问题。

(2) 数据检索:是指以数据为检索对象,其检索结果是可供直接使用的科学数据的检索。数据检索是一种确定性检索,例如:查找“艾滋病毒在体外血液中的存活时间是多少?”就是一种确定的数据检索。

(3) 事实检索:是指以某一客观事实为检索对象,其检索结果主要是客观事实或为说明客观事实而提出的数据的检索。事实检索也是一种确定性检索,例如:查找“在2001年世界上有哪几个国家、哪些实验室在研究炭疽菌?”就是一种确定的事实检索。

但是文献检索、数据检索和事实检索三者并没有严格的界限划分,文献检索这个用语随着社会信息化进程的加快,其内涵也更为丰富和广泛。可以说,根据课题的特定要求,查找出课题所需的特定文献、数据和事实的整个过程都是文献检索的范畴。

3. 按照检索手段划分 按照检索手段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。

(三) 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,它们的实现有赖于检索系统的存在。文献的存储过程就是检索系统的制作过程,这一过程通常由信息工作者去完成;而文献的检索过

程则是检索系统的利用过程,可由文献需求者或情报人员去完成。所以任何检索系统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,就是按照既定的检索系统建设方针、目的、标准等从信息源中选择所需文献,对这些文献的外表特征和内容特征进行分析和加工,形成文献特征的标识,再把这些标识按事先设计好的规则进行排序形成多种索引编入到检索系统中去的过程。

文献的外表特征,也称为自然标识,包括标题、作者、来源、卷期、页次、年月、类型、号码、文种等项目。文献的内容特征,也称为人为标识,是指文献论述的主题,即文献的中心内容,诸如主题词、分类号、类目名称、文摘等。对文献内容特征的存储过程也就是文献标引的过程。文献标引(indexing)是指对原始文献的内容深入理解,进行主题分析,把握住它所论述的中心内容,形成主题概念,然后选用特定的情报检索语言(即词表或分类法)来表达其主题概念,转换成系统语言,形成文献特征标识的过程。如对文献赋予分类号标识的过程称为分类标引,赋予主题词标识的过程称为主题标引。在计算机检索系统中,也就是检索数据库中,每一种文献特征称为“字段”,文献特征标识就是字段值,所有的文献特征(字段)即构成一篇文献的题录或文摘(记录)。

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(即文献特征标识);最后利用这些特征标识到检索系统中去查找到相关文献线索或具体的文献。

由此可知,检索过程是与存贮过程相对应的逆过程,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。而连接这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。

四、文献检索系统

1. 文献检索系统的定义 文献检索系统是指根据特定的文献信息需求而建立起来的一种有关文献信息的搜集、加工、存储和检索的程序化系统。根据医学文献信息存储与检索所采用的设备和手段,医学文献检索系统可分为手工检索系统和计算机检索系统两种类型。近年来,计算机检索发展十分迅速,基本代替手工检索。因此,当今文献检索系统尤指计算机信息检索系统。

2. 文献检索系统的组成要素 文献检索系统通常由以下4个基本要素组成:

(1) 检索文档:检索文档是标有检索标识的信息集合,如书目、索引或文摘等手工检索工具中由文献款目组成的正文、工具书中由条目或短文组成的主体,计算机信息检索系统中的数据库、主文档、倒排文档等。

(2) 技术设备:技术设备是指能存储信息及其标识以及实现存储和检索操作的各种软硬件设备。如手工检索系统的卡片目录或检索刊物,计算机检索系统的输入输出装置、存贮器、控制器、运算器等,网络检索时还需要网络适配器、通信线路、通信装置、终端设备及其相应的软件等。

(3) 语言工具:语言工具指检索语言、标引规则、输入输出标准等。

(4) 人员:主要包括信息加工标引人员、录入人员、检索人员、系统管理维护人员等。

这些人员是检索系统中的主导因素。

3. 文献检索系统的内容结构 手工文献检索系统和计算机文献检索系统的内容结构组成有所不同。

(1) 手工检索系统的内容结构: 手工检索系统一般由五个部分组成: 编辑使用说明、目次表、正文、索引和附表。①编辑使用说明: 主要是为帮助使用者了解和使用该检索工具而编写的, 为使用者提供必要的指导, 一般包括编制方法和原则、使用范围、收录年限、著录格式、查找方法及注意事项, 以及所使用的代号说明等。②目次表: 揭示正文部分或者整本检索刊排列的名称、排列次序及所在页码。③正文: 这一部分为检索工具的核心, 由按一定规则排列的一篇篇文献条目组成; 一般为了方便识别, 每篇文献条目按照前后顺序赋予顺序号, 一个顺序号代表一条文献条目。④索引: 检索工具的检索途径(检索入口)主要体现在其所附的索引部分, 一般一种索引就是一种检索途径。常用的索引有主题索引、著者索引、分类索引、专利索引等。⑤附表: 一般包括所收录的刊物名称, 所使用文字的翻译, 各种简、全称对照, 术语以及所使用的主题词表, 甚至文献入藏单位及代号等。

(2) 计算机文献检索系统的内容结构: 计算机文献检索系统主要由主文档、倒排文档和使用帮助导航3部分组成。为了能够进行随机检索, 每个计算机信息检索系统都存有一个主文档和若干个倒排文档。①主文档: 又称为顺排文档, 它是将文献检索数据库中全部记录按一定顺序排列而成的文献记录集合。一般情况下, 数据库的顺排文档是按记录的流水号(或存取号)的大小从小到大顺序排列的, 相当于手工检索工具的正文部分。②倒排文档: 又称为索引文档, 是将文献检索数据库全部记录中可检索字段及其内容(字段值)提取出来, 再按其字段值的某种顺序排列而成的集合。倒排文档存入了数据库全部记录的文献特征标识(即索引单元)。一般情况下, 倒排文档是按文献特征标识的字母顺序排列的, 每个文献特征标识后面都标有相应记录(文献)的存取号和字段位置标识符, 相当于印刷型检索工具的辅助索引部分。数据库的检索中, 计算机就根据这些存取号和字段位置标识符去识别每个文献特征标识所处的具体记录和字段位置。③帮助导航: 计算机文献检索系统一般在主页或主屏幕上会提供帮助或使用指南等的链接, 有的系统还提供针对当前操作页面可能遇到的问题的即时帮助信息。用户在使用系统过程中如果遇到任何问题都可以寻求帮助或指南进行解决。

4. 文献检索系统的类型 按照所摘录的文献的加工程度, 文献检索系统可分为目录、题录、文摘和全文检索系统4种类型。

(1) 目录型检索系统: 是对图书或其他单独出版的资料的系统化记载及内容揭示, 通常以图书、期刊作为报道单元的文献检索系统。按照职能划分有出版发行目录、馆藏目录、资料来源目录等; 按收录文献种类划分可有图书目录、报刊目录、标准目录、专利目录等; 按物质载体形式可分为卡片目录、书本式目录、机读目录等。

1) 出版发行目录: 包括登记性的国家书目和商业性的出版商书目。国家书目是国家出版物登记制度的产物, 是有关一个国家全部出版物的现状和历史的记录。目前我国的国家书目有两种: 月刊性的《全国新书目》和年刊性的《全国总书目》, 国外较著名的有《英国国家书目》、《法国国家书目》等。

2) 馆藏目录和联合目录: 馆藏目录是反映单个图书馆文献信息收藏情况的目录型检索系统, 例如《北京图书馆馆藏目录》。联合目录则是反映某一地区或系统甚至全国的多个图书馆的文献收藏情况的目录型检索系统, 例如《中国图书联合目录》、《全国西文期刊联合目录》。图书馆之间的馆际互借、资源共享是图书馆发展的大势所趋, 联合目录是实现图书馆

之间馆际互借、资源共享的前提。

3) 资料来源目录:是指附属于某一检索系统或工具的“引用出版物目录”,有的也称“资料来源目录”,例如美国《医学索引》的《收编期刊一览表》。

(2) 题录型检索系统:是以一篇篇内容上独立的文章作基本著录,并只著录文献的外部特征,包括文献篇名、著者、著者单位、原文出处、语种、专利号、合同号等各项内容的文献检索系统。不同类型文献的题录内容不尽相同,因而不同检索系统的题录著录项目也不尽相同。

题录和目录虽然都是对文献外部特征进行描述,但二者是不同的。目录著录一个完整的出版单元,而题录著录的只是一个完整的出版物的某一部分,例如期刊中的论文、图书中的章节等。题录型检索工具的中文名称有“题录”、“索引”、“目录”等,例如《全国报刊索引》、《中文科技资料目录》、《放射医学题录》等,英文名称可有“Index”、“Title”等,例如美国的《Index Medicus》、《Chemical Titles》。

(3) 文摘型检索系统:除了描述文献的外部特征之外,还对文献的内容特征作较深入的介绍,即带有文献的摘要内容的文献检索系统。目前国内外科技论文的撰写都要求附带较高质量的文摘,尤其是外文文献,在很多情况下,甚至不用看原文就能理解文章的大致思想内容。因为文摘型检索系统比题录和目录型检索系统揭示文献内容要深,因此它所报道的速度相对要慢些。

(4) 全文型检索系统:是不仅能获取文献的题录线索信息,还能得到文献的原文内容的文献检索系统。它一般为计算机检索系统,如《中国期刊全文数据库》、《保健、医学与药学电子期刊全文数据库》等。

此外,文献检索系统还有很多种不同的分类方法。如按照收录文献类型不同可分为期刊检索系统、图书检索系统、专利检索系统等;按照载体形式不同可分为书本式检索系统、磁带式检索系统、缩微式检索系统、光盘检索系统、网络检索系统等;还有前述的按加工文献和处理信息的手段不同可分为手工检索系统和计算机检索系统。

5. 文献检索系统的评价 文献检索系统的种类繁多、数量庞大、质量高低不一,判定检索系统质量的因素主要包括以下几个方面:

(1) 报道信息的准确性:亦即质量保证,包括揭示信息外部特征和内容特征的准确程度及检索出来的信息符合实际需求的程度。

(2) 报道信息的及时性:可通过时差、新颖率、更新周期等指标来反映。时差是指原始信息发表或出版到它在检索系统中反映出来之间的时间间隔或时间滞留;新颖率是指用户在检索前没有掌握的信息在被检索出来的信息中的比例;更新周期则为检索工具连续出版的频率。

(3) 索引体系的完善程度:主要是指检索手段是否完善,检索功能和检索途径是否齐全。

(4) 对信息标引的深度:标引深度反映了对信息内容特征进行描述的细致程度,或者说分析信息主题内容所达到的深度。标引深度是决定检索系统质量高低的重要因素。

(5) 查全率和查准率:查全率(Recall Ratio)和查准率(Precision Ratio)是检索系统最为流行和重要的两个性能和效果评价指标。查全率(R)和查准率(P)的定义可用如下公式表示:

$$R = \frac{\text{检出的相关信息量}}{\text{检索工具中相关信息总量}} \times 100\%$$

$$P = \frac{\text{检出的相关信息量}}{\text{检出的信息总量}} \times 100\%$$

需要注意的是,检索系统中针对某一提问的全部相关信息数量不能精确获知,因此,R值的计算一般为近似值。另外R和P之间具有密切关系,英国情报学家克莱弗登(C. M. Cleverdon)在其著名的Cranfield II试验中发现R和P之间存在互逆关系,即如果检索工具的查全率较高,则其查准率将相对下降;反之查准率高,则查全率低;而且查全率和查准率只能相对提高,二者永远不可能同时达到100%。

总之,影响检索系统质量高低的因素很多。对用户而言,要检索出满意的结果,选择质量高的检索系统是其中一个重要环节,此外用户也要提高自身利用检索工具的水平,如合理选择检索点、适当调整对查全率和查准率的要求等,以便最大限度地发挥检索工具的作用。

五、文献检索的意义

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“继电器接点电路合成研究”,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,了解当前其发展水平,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于一定的方法,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。有人说,20世纪70年代的大学本科毕业生就业后其知识可以用8年,80年代可以用5年,而90年代只能用两年。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