

李勇文·主编

医学文献 查询与利用

YIXUE
WENXIAN
CHAXUN
YU
LIYONG



四川大学出版社

医学文献 查询与利用

YIXUE WENXIAN
CHAXUN YU LIYONG

主 编 • 李勇文

副主编 • 龙兴跃 伍 利 刘 娟

编 者 • 曾满江 胡 臻 丘 琦

冉 黎 张 容 夏 莹

吴丽娟



四川大学出版社

责任编辑:梁 平
责任校对:田凤桢
封面设计:璞信文化
责任印制:王 炜

图书在版编目(CIP)数据

医学文献查询与利用 / 李勇文主编. —成都:四川大学出版社, 2017. 2

ISBN 978-7-5690-0371-0

I. ①医… II. ①李… III. ①医学文献—情报检索
IV. ①G257.36

中国版本图书馆 CIP 数据核字 (2017) 第 036518 号

书名 医学文献查询与利用

主 编	李勇文
出 版	四川大学出版社
地 址	成都市一环路南一段 24 号 (610065)
发 行	四川大学出版社
书 号	ISBN 978-7-5690-0371-0
印 刷	郫县犀浦印刷厂
成品尺寸	185 mm×260 mm
印 张	19.25
字 数	470 千字
版 次	2017 年 2 月第 1 版
印 次	2017 年 2 月第 1 次印刷
定 价	42.00 元



版权所有◆侵权必究

◆读者邮购本书,请与本社发行科联系。

电话:(028)85408408/(028)85401670/

(028)85408023 邮政编码:610065

◆本社图书如有印装质量问题,请
寄回出版社调换。

◆网址:<http://www.scupress.net>

前 言

在人类社会已进入信息时代的今天,信息资源在日常生活和经济社会发展中扮演着越来越重要的角色。其中,文献信息资源是当今社会人们学习、工作和生活不可或缺的重要组成部分。

以人类的疾病谱及健康观念的变化为依据,医学模式由传统的生物医学模式向生物—心理—社会—环境医学模式转变,新医学模式的研究对象不仅是自然人,还要研究人的状态和所处的环境。这种模式的转变,对医学生的信息素养提出了更高要求,要求医学生必须具有敏锐的信息意识、娴熟的信息能力和良好的信息道德。

《医学文献查询与利用》主要面向高等医学院校学生,是一门关于如何查询、获取和利用医学文献信息的工具性课程,旨在培养医学生的信息素养和独立学习的能力。当学生初入医学院校大门时,面对图书馆的海量馆藏文献资源,他们对于如何快速有效找到自己所需要的知识往往束手无策,《医学文献查询与利用》就是他们开启知识宝库的金钥匙,是他们通往成才之路的好帮手。

本教材以原教育部于1985年9月印发的《关于改进和发展文献课教学的几点意见》和原国家教委于1992年5月印发的《文献检索课教学基本要求》为基本指导,结合现代信息技术发展趋势以及《全球医学教育最低基本要求》中对医学生的信息能力要求,强调以下编写原则:以“实用、够用”为原则,以医学院校学生的自主学习和临床医疗信息需求为出发点设计内容体系,深浅、详略得当;注重学生自主学习、协同学习和协同科研能力的培养;基础知识和检索技能相结合,以提升检索技能为主;便于医学文献检索课程教学的实施,既是课堂教材,也是自主学习和开展科研活动的重要参考工具书;根据不同的专业和学历层次可对内容进行灵活调整;注重培养学生的综合信息素质;适应信息技术飞速发展的客观实际,内容体系具有前瞻性。

本教材共10章,第1章概述文献信息资源和信息素养,以及检索原理、技术与策略;第2章讲解如何查找和利用馆藏文献资源;第3章主要讲解各类网络医学信息资源和搜索引擎;第4~8章,分别介绍中文医学文献检索工具、外文医学文献检索工具、引文数据库、特种文献数据库以及循证医学信息的检索与利用;第9章的主要内容为综合检索及各学科的检索案例分析;第10章的主要内容为医学文献评价与利用。

李勇文设计本教材内容体系架构,完成60%以上内容编写,并对全书进行统稿;龙兴跃参编第2章、第3章和第10章;伍利参编第5章和第9章;刘娟参编第6章和第9章;曾满江参编第3章、第10章;胡臻参编第2章和第7章;丘琦参编第3章、第9章和第10章;冉黎参编第5章、第7章、第8章和第9章;张容参编第4章;夏莹参编第4章;吴丽娟参编第7章。

本教材在编写过程中得到了成都医学院卯辉副校长的多方面指导。四川大学出版社为该教材的审稿、出版提供了帮助。在此向卯辉副校长、各位参编人员、被引文献作者、被引网站等表示衷心的感谢！

囿于编者的学识与水平，书中难免有不足之处，敬请读者批评指正！

编 者

2016 年 12 月

目 录

第一章 绪论	(1)
第一节 医学教育与信息素养	(1)
第二节 文献信息资源的类型与特点	(3)
第三节 信息检索原理	(8)
第四节 检索技术与策略	(14)
第二章 图书馆信息资源利用	(22)
第一节 馆藏目录查询系统 (OPAC)	(22)
第二节 数据库导航	(27)
第三节 移动图书馆	(29)
第三章 网络医学信息资源	(33)
第一节 网络医学信息资源概述	(33)
第二节 搜索引擎	(36)
第三节 开放获取资源	(48)
第四节 医学站点	(55)
第四章 中文文献检索工具	(65)
第一节 中国知网 (CNKI)	(65)
第二节 维普期刊资源整合服务平台 (VIP)	(86)
第三节 万方数据知识服务平台	(92)
第四节 中国生物医学文献服务系统	(100)
第五节 中文资源整合检索平台	(109)
第五章 外文文献检索工具	(122)
第一节 MeSH 与 PubMed	(122)
第二节 Elsevier 数据库	(138)
第三节 其他外文数据库	(145)
第六章 引文检索	(154)
第一节 引文检索概述	(154)
第二节 Web of Science	(155)
第三节 中文引文检索	(164)
第七章 专类文献检索	(168)
第一节 专利文献检索	(168)
第二节 学位论文检索	(175)

第三节 会议文献检索·····	(180)
第四节 标准文献检索·····	(187)
第五节 生物信息检索·····	(193)
第八章 循证医学检索 ·····	(208)
第一节 循证医学基础知识·····	(208)
第二节 循证医学的检索资源·····	(211)
第三节 基于 PICOS 原则的循证医学检索 ·····	(215)
第九章 检索实例分析 ·····	(220)
第一节 综合检索分析·····	(220)
第二节 学科案例分析·····	(223)
第十章 医学文献管理与利用 ·····	(243)
第一节 文献管理软件·····	(243)
第二节 医学科研论文写作·····	(254)
第三节 科研评价·····	(260)
第四节 信息道德与学术规范·····	(267)
附录一 《中国图书馆分类法》(第五版) R 类简表 ·····	(275)
附录二 MeSH 范畴表主要类目 (2016) ·····	(280)
附录三 MeSH 副主题词等级表 (2016) ·····	(284)
附录四 副主题词适用范围说明 (2016) ·····	(288)
附录五 ICD-10 常用编码表 (Version: 2016) ·····	(293)

第一章 绪论

第一节 医学教育与信息素养

一、信息素养的概念与内涵

信息素养 (Information Literacy) 是人们能够判断确定何时需要信息, 并能够对信息进行检索、评价和有效利用的能力。信息素养包括信息意识、信息技能和信息伦理三方面。

信息意识指人的信息敏感程度, 是人们对自然界和社会的各种现象、行为、理论观点等, 从信息角度的理解、感受和评价。通俗地讲, 面对不懂的东西, 能积极主动地去寻找答案, 并知道到哪里、用什么方法去寻求答案, 这就是信息意识。信息时代处处蕴藏着各种信息, 能否很好地利用现有信息资料, 是人们信息意识强不强的重要体现。使用信息技术解决工作和生活中问题的意识, 是信息素养教育中最重要的一点。

信息能力是指能够有效地获取、加工和利用信息的能力, 包括信息系统的基本操作能力, 信息的采集、传输、加工处理和应用的能力, 以及对信息系统与信息进行评价的能力等。这也是信息时代重要的生存能力。身处信息时代, 如果只是具有强烈的信息意识和丰富的信息常识, 而不具备较高的信息能力, 还是无法有效地利用各种信息工具去搜集、获取、传递、加工、处理有价值的信息, 不能提高学习效率和质量, 无法适应信息时代对未来医务工作者的要求。信息能力是信息素质诸要素中的核心, 医学生必须具备较强的信息能力, 否则难以在信息社会中生存和健康发展。

信息伦理是指个人在信息活动中的道德情操, 能够合法、合情、合理地利用信息解决个人和社会所关心的问题, 使信息产生合理的价值。培养学生具有正确的信息伦理道德修养, 要让学生学会对媒体信息进行判断和选择, 自觉地选择对学习、生活有用的内容, 自觉抵制不健康的内容, 不组织和参与非法活动, 不利用计算机网络从事危害他人信息系统和网络安全、侵犯他人合法权益的活动。

二、医学文献检索课程与信息素养

彼德·德鲁克在《后资本主义社会》中说: “对于任何一个人、组织、企业 and 国家, 获取和应用知识的能力是竞争成败的关键。” 也就是说, 在现代信息数字化时代, 生存的基本技能是终身学习的技能, 即具有较强的信息发现、吸收、创新能力, 信息素质的

提高成为信息社会的根本。信息素质的提高既包含个体的提高也包含群体的提高。在知识经济时代,学习观念所发生的这种根本改变对学校教育提出了更高的要求,同时也对个人的学习提出了挑战。为了有效解决信息总量剧增与个人学习能力有限之间的矛盾,信息数量激增与信息质量无保证之间的矛盾,网络共享的公平和开放原则与信息壁垒、数字鸿沟之间的矛盾,必须依靠人们信息素养的提高。

信息素养的提高不是一蹴而就的事情,必须接受专业的教育和经过专门的训练才能实现。英国作家、批评家赛缪尔·约翰逊(Samuel Johnson)曾将知识分成两类:一类是我们要掌握的学科知识,另一类是要知道在哪儿可以找到有关信息的信息。对信息素养的培养首先起始于我们对于“知道在哪儿可以找到有关信息的信息”能力的培养,即检索能力的培养。开设医学文献检索课程的目的就在于在适当的时候将适当的信息传递给适当的人,让读者知道“在哪儿可以找到有关信息的信息”,同时讲述检索方法与技巧,帮助读者循序渐进地培养自己发现、吸收、整理、评价与重组信息的能力。

对于医学生而言,在校学习期间不可能学到将来工作中所需的全部知识,因此,最重要的是培养独立获取知识的能力。文献检索就是培养学生这种能力的一门重要课程。通过学习掌握文献检索的基本知识和方法,学生不仅能在浩瀚的知识海洋里准确、迅速、全面、系统地找到所需要的文献资料,而且能掌握自我学习的方法和技术,提高终身学习的能力。这对于医学生增强信息意识和形成合理的知识结构、提高文化素养和专业本领,都具有深远的意义。

三、医学生信息素养能力标准

1999年6月9日,受美国纽约中华医学基金会(简称CMB)资助,国际医学教育组织(Institute for International Medical Education,以下简称IIME)在纽约成立,其主要工作是在定义“全球医学教育最基本要求”方面发挥领导作用。通过“最基本要求”,使在任何国家培养的医生都能达到在医学知识、技能、职业态度、行为和价值观等方面的最基本要求。

GMER是Global minimum essential requirements in medical education的缩写,中文意思为“全球医学教育最基本要求”。IIME(国际医学教育组织)将“最基本要求”归纳为7个领域和具体的60条标准。其中第6大项为“信息管理”能力,指出医疗实践和卫生系统的管理有赖于有效的源源不断的知识和信息,计算机和通信技术的进步对教育和信息的分析和信息管理提供了有效的工具和手段,使用计算机系统有助于从文献中寻找信息,分析和联系病人的资料,因此,毕业生必须了解信息技术和知识的用途和局限性,并能够在解决医疗问题和决策中合理应用这些技术。本项设5条标准:从不同的数据库和数据源中检索、收集、组织和分析有关卫生和生物医学信息;从临床医学数据库中检索特定病人的信息;运用信息和通信技术帮助诊断、治疗和预防,以及对健康状况的调查和监控;懂得信息技术的运用及其局限性;保存医疗工作的记录,以便进行分析和改进。

2000年,ACRL(美国大学与研究图书馆协会)通过的“美国高等教育信息素养能力标准”(Information Literacy Competency Standards for Higher Education),是世界

上影响最大的信息素养标准,为全球高等教育提供了讨论信息素养的概念框架,被世界各国广泛采纳使用。该标准描述了具备较高信息素养的个体应具备的能力,包括5个指标,并更进一步细分为22个能力指标,具有很强的可操作性。标准的引言部分介绍了信息素养的概念并阐述了信息素养与信息技术、高等教育、标准利用以及评估之间的关系。

“美国高等教育信息素养能力标准”的5个大指标分别为:有信息素养的学生有能力决定所需信息的性质和范围;有信息素养的学生可有效地获取所需信息;有信息素养的学生能评价信息及其来源,并能把所选择的信息与原有的知识背景和评价系统结合起来;有信息素养的学生无论是个体还是团队的一员,都能有效利用信息实现特定的目标;有信息素养的学生懂得有关信息技术的使用所产生的经济、法律和社会问题,并能合理合法地获取信息。

教育部、卫生部于2008年9月印发了《本科医学教育标准:临床医学专业(试行)》。在基本要求中,对医学生的信息素养提出了明确要求。在思想道德与职业素质目标部分,提出了“树立终身学习观念,认识到持续自我完善的重要性,不断追求卓越”的信息素养要求;在技能目标部分,提出了“运用循证医学的原理,针对临床问题进行查证、用证的初步能力”“结合临床实际,能够独立利用图书资料 and 现代信息技术研究医学问题及获取新知识 with 相关信息”“具有自主学习和终身学习的能力”等信息素养要求。

在当今的信息社会,以计算机和网络为核心的信息技术已逐渐渗透到社会的各个层面,深刻地影响和改变着人们的生活,信息和信息技术日益成为社会各个领域中最活跃、最具有决定意义的因素。医药卫生人才结构需求的重大变革对传统的高等医学教育标准和医学生信息素养提出了新的挑战。信息素养作为一种获取、评价和利用信息资源解决问题的能力,也是临床医疗工作的必要条件和必备素质。对于医学生与医务工作者而言,信息素养不仅指平时必须具备的关于信息方面的一般观念、意识、知识与技能,还包括在从事医疗工作中所具备的信息处理技能,对信息进行筛选、鉴别和利用的能力,以及进行医学科学研究所应具备的获取、分析、利用和创造信息的综合能力。

第二节 文献信息资源的类型与特点

一、文献信息资源类型的划分

1. 按载体进行划分

(1) 书写型。

该类型一般以纸张或竹简为载体,用人工抄写而成,如手稿、书法作品、医生写的病案记录、各种原始记录和档案等。

(2) 印刷型。

该类型是指以纸张为载体印刷而成的文献,目前仍是出版物的主要形式,也是馆藏

文献的主要类型,如图书、期刊、特种文献等。

(3) 缩微型。

该类型以感光材料为载体,用摄影的方法把文献记录在胶卷或胶片上,如缩微胶卷、缩微平片等。

(4) 视听型。

该类型即视听资料或声像资料,包括唱片、录音带、录像带、电影片、幻灯片等。这种文献直接记录声音和图像,如心脏病变的杂音、外科手术过程,可给人以具体的视觉形象和听觉感受,犹如身临其境。

(5) 机读型。

该类型是指以数字化技术将文献贮存在磁带、磁盘、磁鼓或光盘上,通过计算机阅读的文献。这种文献的存储容量大,检索速度快,使用方便,在文献检索和全文存储方面显示出特殊的优越性,如电子图书、电子期刊、文献数据库等。

2. 按出版形式进行划分

(1) 图书。

图书是现代印刷出版最常见的一种,内容广泛,系统地论述一个专题,是掌握一门学科的基本资料。在每一种正式出版的图书的版权页或其他明显部位都有一个由13位(2007年1月1日前为10位)数字或字母组成的国际标准书号(ISBN),这是一种国际通用的出版物代码,代表某种特定图书的某一版本,具有唯一性和专指性,读者可借此通过某些文献信息系统查询某种特定图书。图书根据其内容、作用可分为一般性图书和工具书。

①一般性图书是图书馆的主要藏书之一,其共性是全面、系统地论述某一方面或专题内容的文献。常见的图书有:

A. 教科书及其教学参考书(Textbook),反映本学科的基本知识,是教学的基本用书,其内容比报刊成熟、定型。

B. 图谱(Atlas),是学习有关学科知识重要的参考书,如《正常人体解剖学图谱》。

C. 专著(Monograph),内容精、深,专业性强,往往是科学研究某一课题的总结或某一领域的历史发展、成果内容的集合。

D. 著作集或选集(Selection),是为纪念某名人或著名科学家,出版其生平所著的论文或记录其科学成就。

E. 丛书(Series),是成套的图书,按专题分册单独出版或成套发行。

②工具书(Reference Book),是广泛收集某一范围的知识或资料,并按特定体例或方式编排,能提供资料或资料线索而不是供系统阅读的图书。其特点是内容广泛、信息量大、概括性强、可信度高、便于检索。工具书主要有:

A. 字典、词典(Dictionary),是主要用于解释字的形、音、义,事物及术语的工具书,如《汉语大词典》《实用医学词典》等。

B. 百科全书(Encyclopedia),是综合性工具书,通过收集自然科学、社会科学、科学史以及名人传记等而成,按学科分册,一般卷册数多,从几册到上百册,如《中国

大百科全书》《中国医学百科全书》等。

C. 年鉴 (Yearbook), 是概括评述一年中某学科或分支学科资料的参考书, 每年出版一次, 如《中国百科年鉴》《中国卫生年鉴》等。

D. 手册 (Handbook), 是汇编某一学科专业的基础知识、基本资料或数据供读者查阅用的工具书, 如诊疗手册, 内容包括常见病诊断和治疗原则、临床检验正常数值、常规操作等方法等。

E. 指南 (Guide), 为一般性工具书, 有参考图表、科技数据、工作方法、步骤、过程等。

F. 目录 (Bibliography), 又称书目, 是以文献的自然出版形式为单位来记录文献, 只供检索, 主要报道实用的文献或收藏文献的情况, 如《全国总书目》《科技新书目》等。

(2) 期刊。

期刊是定期或不定期的连续出版物, 有固定的刊名和出版形式, 有年、卷、期号。期刊具有内容新颖、出版周期短、刊载论文速度快、品种多、涉及面广等特点, 能及时反映科技水平、科研动态, 是科技情报的主要来源。核心期刊, 指的是刊载与某一学科(或专业)有关的信息较多且水平较高、能够反映该学科最新成果和前沿动态、受到该专业读者特别关心的那些期刊。同图书的 ISBN 号一样, 每种期刊均有一个由 8 位数字组成的国际标准连续出版物号 (ISSN), ISSN 同样具有唯一性和专指性, 因而成为读者查询某种刊物的一个检索途径。期刊种类众多, 主要有:

A. 专业期刊/杂志 (Journal), 有专业性、综合性的杂志, 医学各学科的杂志属于医学专业杂志, 如《中华医学检验杂志》。

B. 学报 (Acta), 是水平较高的科学杂志, 由专业学会或高等院校出版, 主要刊登学科的原始学术论文, 如《第三军医大学学报》。

C. 通报 (Bulletin), 是综合报道性期刊, 主要报道各学科的现状, 如《中国药理学通报》。

D. 综述或述评 (Review), 是对某一专题进行综合概括、深入评论叙述的期刊, 如《眼科新进展》。

E. 文摘 (Abstract), 是用文摘形式报道的期刊, 如《中国医学文摘》系列和《健康文摘》等。

F. 索引 (Index), 是以题录形式报道的期刊, 如美国《医学索引》(Index Medicus, IM) 和《传染病专题索引》等。

(3) 特种文献。

特种文献为非书非刊的文献, 包括专利文献、科技报告、学位论文、标准文献、技术档案、政府出版物、会议录等, 举例如下:

A. 专利文献 (Patent Document), 是由国家专利局公布或正式归档的与专利有关的文献, 包括专利说明书、专利公报、专利分类表、专利索引以及从专利申请到批准全过程的一切文件和资料等。

B. 科技报告 (Scientific & Technical Report), 是某项科研项目提出的正式报告或

进展情况的报告,内容具体、专深,反映新的科研课题和高科技方面的信息。

C. 学位论文 (Dissertation),是学术界培养的博士、硕士生,通过科学研究、实验研究及论文答辩,取得学位资格的论文。这些论文有很好的参考价值,但大多不公开出版发行,属图书馆特藏。

D. 标准文献 (Standard Document),是指对产品或工程质量所作的技术规定,具有一定法律效力,很多标准是从事生产建设和科研工作的依据。

E. 技术档案 (Technical Archives),是科技工作中形成的技术性文献,如科研规划、设计方案、工程图表、实验记录、病案记录等。这类资料由专业人员整理,可靠性强,具有较大的使用价值。

F. 政府出版物 (Government Publication),是指各国政府及其所属机构出版的文献资料,内容广泛,有行政和科技之分,包括政府法规、方针政策、统计资料等。

G. 会议录 (Proceeding),是综合报道学术会议讨论交流的论文、报告及情况的一种出版物。

二、文献的级别

医学文献根据其内容、结构、性质及加工程度不同可分为一次文献、二次文献、三次文献和零次文献四个级别。

(1) 一次文献 (Primary Document),一次文献又称原始文献,是作者以生产和科研成果为依据而创作的原始论文。其特点是含有前所未有的发明创造,或者新的见解和理论。专题著述、期刊论文、科技报告、专利文献、学位论文、会议资料等均属一次文献,是科技文献的主体。但由于其量大、分散而无序,给读者的查找与利用带来极大的不便。

(2) 二次文献 (Secondary Document),是将分散无组织的一次文献进行收集、整理、压缩、加工,并按一定的顺序组织编排而成的检索工具,包括目录、题录、文摘、索引等。二次文献通常由图书情报机构组织编辑出版,是对一次文献进行加工,如著录文献特征、摘录内容要点、标引文献主题、按学科进行分类等,使之成为有组织、有系统的检索工具。

从上述定义的引申来看,二次文献信息是关于文献的文献、关于信息的信息。因此,现在网上的百度、谷歌等主题指南、搜索引擎,都是关于数据库和网页的信息集合,其功能作用等同于上述二次文献,所以称其为网络检索工具。相对于一次文献而言,二次文献是从分散到集中、从无序到有序、从繁杂到简约,因而具备了可查检的便捷性,用以解决读者查阅所需特定文献线索的问题。正因为如此,包括网上检索工具在内的二次文献及其利用也就成为文献信息检索的核心内容。

(3) 三次文献 (Tertiary Document),是对一、二次文献进行综合分析研究,作出系统整理和概括的论述文献。三次文献是对知识、情报的第三次加工,是利用二次文献收集大量相关一次文献,对其内容进行分析综合、重新组织加工而成。属于三次文献的有综述、述评、进展、现状、发展趋势等期刊文献和百科全书、年鉴、手册等参考工具书。三次文献具有信息量大、综合性强和参考价值大等特点,可使读者不必大量阅读一

次文献,就可比较全面地了解某一专题、某一领域当前的研究水平、动态。

(4) 零次文献 (Zero Document), 指未经信息加工, 直接记录在载体上的原始信息, 如实验数据、观测记录、调查材料等。这些未融入正式交流渠道的信息, 往往反映的是研究工作取得的最新发现, 或是遇到的最新问题, 或是针对某些问题的最新想法等, 而这一切无疑是启发科研人员的思路, 形成创造性思维的最佳思维素材。

此外, 学术界还常将通过非正式交流渠道获得的非正式出版物称作灰色文献 (Grey Literature)。灰色文献和零次文献的概念内涵有一定程度的重叠, 但作为一般的专业人员可不必严格区分这两个概念。

三、医学文献信息的特点

随着生命科学世纪的到来, 医学科学技术迅速发展, 作为记录医学信息知识的载体和保存、传播医学知识的医学文献, 表现出以下发展特征和趋势。

1. 文献数量庞大, 增长速度加快

科技文献中, 医学及生物科学文献总量非常庞大, 增长速度占各学科之首。以占文献总量 70% 的期刊为例, 全世界期刊总数有 14 万余种, 其中生物医学期刊已超过 21000 种, 约占 1/7。美国《科学引文索引》(SCI) 按引文数量排列的前 500 种期刊中, 医学期刊 176 种, 约占 1/3。其他各类文献中, 医学类文献所占比例基本相似。

2. 学科交叉渗透, 内容分散重复

现代科学技术发展的一种趋势是科学门类高度分化又高度综合, 新的分支和边缘交叉学科不断产生, 学科之间互相渗透, 致使文献分散, 在内容与结构上又产生交叉。医学论文不仅刊载在医学专门期刊上, 还大量地刊载在一些综合性期刊和其他相关科学领域的期刊上。近年来, 由于多种因素的影响, 文献重复发表的现象屡见不鲜, 出现一文多刊、转载互译等现象。如此分散、交叉、重复, 明显增加了查寻搜集文献的难度, 而且大大增加了文献量的冗余。

3. 知识更新频繁, 文献发表滞后

科学技术的发展, 对科学奥秘的探索和认识不断深化, 知识更新愈来愈频繁, 导致记录知识的文献老化速度加快。19 世纪老化速率为 50 年左右, 而如今已缩短到 5~10 年。国外研究发现, 生物医学文献的半衰期为 3 年, 物理学为 4.6 年, 化学为 8.1 年, 植物学为 10 年, 数学为 10.5 年, 地理学为 16 年, 可见生物医学文献的老化速率较快。此外, 医学文献的发表速度比医学文献增长的速度要慢得多, 原因是大量的文献不能及时发表。论文从编辑部收到稿件至正式发表可长达一两年之久, 使得一些文献正式发表时已失去了某些应有的价值。为此, 科技人员之间往往采取直接交换手稿复本、参观访问、会议交流等有效途径获取未发表的文献。

4. 语种不断增多, 交流传播加速

20 世纪初, 只要掌握英、德、法三种语言, 就可阅读全世界 92% 以上的科技文献, 如今全世界的科技期刊涉及的文种已达 70~80 种之多。医学文献涉及的文种也相当多, 如 PubMed 报道了 55 种语言的文献。文种增加, 造成了读者阅读文献的障碍, 影响了

信息情报的交流与传递。

由于现代交通、通讯和印刷技术的发展,情报信息载体的磁性化、机读化以及多媒体和国际互联网的广泛应用,为情报信息的快速传递和交流提供了非常便利的条件,医学文献信息的用户,可以通过互联网在瞬间获取所需要的文献。

5. 文献信息向数字化方向发展

随着计算机、数据存储、数据传输以及通信技术的发展,文献信息由传统的纸质印刷向电子化、网络化、数字化方向发展。20世纪60年代,美国国立医学图书馆首创“医学文献分析和检索系统”(MEDLARS)。1989年中国医学科学院医学信息研究所建成《中国生物医学文献数据库》,向全国医学相关机构提供光盘数据检索服务。

Internet是将各个国家、各个部门、各个领域的不同信息、资源连成一个整体的超级信息资源网络,用户可以通过各种信息查询工具访问所有的信息资源。1996年美国国立医学图书馆免费开放该馆的MEDLINE(即PubMed)等15个数据库。我国2000年4月正式启动“中国数字图书馆”工程,将浩如烟海的、各种形式的文献资料加以数字化处理,并使之流通于全球信息网络。它与Internet上的网页信息资源也有区别,即是经过分类、编辑、整理、加工而成的有序文献资源。目前,数字资源已成为读者利用文献信息资源的主要类型。

第三节 信息检索原理

一、信息检索概述

信息检索是指利用一定的检索算法,借助于特定的检索工具,并针对用户的检索需求,从结构化或非结构化的数据中获取有用信息的过程。我们可以把整个信息检索过程划分为三个方面:信息的存储与组织、信息的检索、信息的展示。图1-3-1给出了信息检索三个方面衔接的原理示意图。

最早的计算机情报检索的试验是美国海军兵器中心于1954年完成的,它主要将文献号以及文献的索引词输入计算机,检索采用对索引词检索获取文献号。这虽然是一个试验性的项目,但它无疑开创了计算机信息检索的先河。20世纪90年代以来,特别是Internet的发展,人们所面对的检索对象更加复杂,检索的需求更加强烈,能够获得的信息类型也是丰富多彩,人们解决各类问题都希望在数据库或互联网中寻找答案。

信息检索对象包括文献、数据、事实等,在学习、工作乃至日常生活中,凡同信息需求有关的问题,均可通过检索各类信息系统来达到获取相应信息的目的。检索方式一般分为手工检索和计算机检索。计算机检索经历了联机检索、光盘检索、网络检索几个阶段,现在人们基本都通过网络检索来获取相应信息。

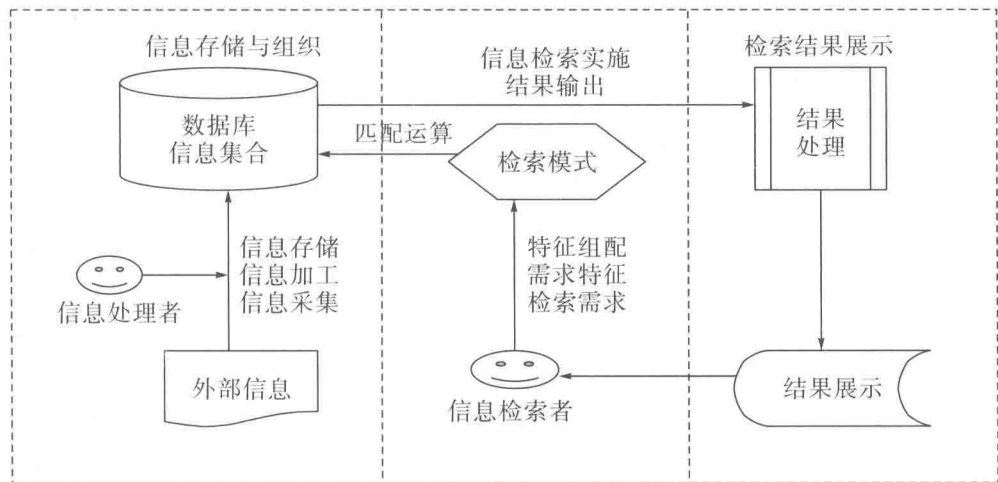


图 1-3-1 信息检索原理示意图

二、信息的组织

从图 1-3-1 可以看出，信息检索全过程包括信息存储组织与信息检索两个过程，信息存储组织是基础，信息组织得越科学、规范，信息检索的质量和效率就越高。信息组织是一种普遍的社会行为，是一切人类活动有条不紊地运行的前提。信息组织是将无序的信息按其外部特征和内容特征有序化，然后进行重新控制，其目的在于提供可控性的高效信息服务。

分类法、主题法以及书目控制是传统信息组织的重要方法，网络环境下，传统分类法既受到挑战，也面临着新的发展机遇。传统分类法的知识系统性和标识语言的通用性及其族性检索能力，是其他情报检索语言所不具备的，也是无法取代的，因此，它在新的信息环境下也仍然得到广泛使用。使用分类法组织网络分类目录并提供浏览方式进行查询，具有直观、信息质量高等优点，也为信息的浏览检索提供了技术保障。用主题词组织与揭示信息具有直接和直观的特点，在网络环境下有相当一部分网络资源浏览器与引擎都以主题词为组织与揭示信息的重要用途与方法。由此可见，分类法与主题法是各种网络信息资源最基本、最主要的组织法与检索法。在下一部分的信息检索语言部分，将对分类法和主题法进行更详细的介绍。

目录组织向来都是信息组织的重要组成部分，在网络环境下，其重要性仍然是非常突出的。人们普遍认为网络信息资源急需像传统文献资源一样进行编目，并提供目录检索。目前网络资源编目模式分为介入与不完全介入两种。前者完全由编目人员提供对信息源的书目描述数据，后者的描述数据可由信息提供者、信息源的管理维护人员等多种途径提供。也就是说，一种方法是由编目工作者进行书目控制，另一种方法则是由信息提供者在提供信息的同时提供信息的描述数据。前者主要以 USMARC 格式的研究为主体，而后者则主要以元数据的研究为主。

信息组织技术包括搜索引擎、元数据、数据挖掘、数据仓库、知识发现、标记语言、数据库技术、自动跟踪技术、机器翻译技术、信息检索的推拉技术、虚拟图书馆技

术、专业指引库技术、智能代理技术、多媒体技术、人工智能技术等。搜索引擎作为网络环境下重要的信息组织工具,自它出现之日起就备受人们的关注,包括图书情报界和计算机界以及信息产业界在内的众多学者对此展开了深入而有效的研究。我们将在第三章给大家讲解搜索引擎在信息检索中的应用技术。

三、信息检索语言

检索语言是用于描述信息系统中信息的内容特征或外表特征和表达用户信息提问的专门语言,是人与信息系统对话的媒介。在各种信息检索系统中,信息都必须被标引或赋予某种简洁的含义,大多数还具有唯一性、科学性、实用性及一定权威性和标准化的标识或编码,既便于计算机处理,又有利于信息的广泛交流与共享。

虽然检索语言主要是信息专业工作者在自然语言基础上创建并使用的,但是由于检索语言是检索者与检索系统之间达到共同理解的基础,检索者有必要学习其中的主要规则、基本原理,才能达到理解一致,减少漏检和误检,提高检索效率。同时,许多医学信息系统,特别是临床信息系统的建设过程中还离不开临床医生的密切合作和参与。

1. 信息检索语言的种类

从表现形式上看,信息检索语言就是文献信息检索系统中的标识系统,能提供多种多样的检索点,如著者、分类号、主题词、关键词等。信息检索语言在各种文献检索系统中无处不在,它种类繁多,各具特点,各有优势又或多或少存在缺陷。在实际应用中常有两种或多种检索语言用于同一检索系统以供选择使用或者相互取长补短。近年来,在强大的计算机信息技术支持下研制和开发的新型检索语言集成系统,已使网络文献信息智能化检索初见端倪,用户的检索提问可以用短语甚至句子等自然语言形式输入,系统能够进行自动分析形成检索策略并进行检索。检索技术的进步,很大程度上得益于检索语言研究成果的应用。

(1) 文献外表特征检索语言。

这是依据文献的外表特征,如文献题名、著者等作为标识和检索点而设计的检索语言。

①文献题名索引系统,以书名、刊名等作为标识的字顺索引系统,如书名目录、引用期刊一览表等;

②以文献上署名的个人作者、译者、编者或学术团体名称作为标识的字顺索引系统,如著者索引、专利权人索引等;

③文献序号索引系统,以文献特有序号为标识的索引系统,如专利号索引、技术标准号索引等;

④引文索引系统,这是利用科学文献末尾所附引用文献、参考文献目录,揭示科学论文系统之间引证和被引证关系而编制的索引系统,如SCI、SSCI、CSCI、CSSCI等,第六章第一节将对此作详细介绍。

(2) 文献内容特征检索语言。

描述文献内容特征的检索语言主要有分类检索语言和主题描述语言两大类。