



中国科学院教材建设专家委员会规划教材
全国高等医药院校规划教材

实用生物医学信息检索

陈 平 张轶群 主编



科学出版社

中国科学院教材建设专家委员会规划教材

全国高等医药院校规划教材

实用生物医学信息检索

主 编 陈 平 张轶群

副主编 陈善志 靳令经 陈 苏 顾 华

编 委 (按姓氏笔画排序)

陈 平 陈 苏 陈善志 张轶群

顾 华 靳令经

科学出版社

北 京

· 版权所有 侵权必究 ·

举报电话:010-64030229;010-64034315;13501151303(打假办)

内 容 简 介

《实用生物医学信息检索》以实用为特点,重点介绍与实际需求密切相关、满足日常学习和工作的生物医学信息检索数据库和网络资源。除详细介绍常见的中、外文医学检索工具基本检索功能外,本书突出笔墨,以图文并茂的形式,着重解释维普、万方、中国知网、InCites、Endnote、HistCite 在信息分析和评价中的作用和功能,使信息检索课程从一般功能介绍向信息分析和评价深度转变,提升和拓展信息检索课程存在的意义和价值。

本书结构合理、条理清晰,适合医学院各专业、生物信息学专业学生作为教材使用,也适合继续教育者、兴趣爱好者自学。

图书在版编目(CIP)数据

实用生物医学信息检索 / 陈平, 张轶群主编. —北京: 科学出版社, 2015. 3
中国科学院教材建设专家委员会规划教材 · 全国高等医药院校规划教材
ISBN 978-7-03-043609-2

I. ①实… II. ①陈… ②张… III. ①生物医学工程—情报检索—医学院校—教材 IV. ①G252.7

中国版本图书馆 CIP 数据核字(2015)第 044982 号

责任编辑: 胡治国 杨鹏远 / 责任校对: 刘亚琦

责任印制: 李 利 / 封面设计: 范璧合

版权所有, 违者必究。未经本社许可, 数字图书馆不得使用

科学出版社 出版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

三河市骏杰印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2015 年 3 月第 一 版 开本: 787×1092 1/16

2015 年 3 月第一次印刷 印张: 13 1/4

字数: 312 000

定价: 35.00 元

(如有印装质量问题, 我社负责调换)

前 言

数字化、网络化、信息化的社会要求人们快速、便捷地从海量信息中找到所需要的有用信息,掌握信息检索基础知识、利用计算机网络检索信息资源是适应时代发展的必然要求。

本书在研究借鉴同类教材和专著的长处、吸收多年信息检索教学实践经验的基础上,结合高等院校对人才培养目标的要求以及课程自身的特点编写而成。全书坚持实用原则,侧重理论联系实际,层次清晰,深入浅出,有较强的针对性和实用性。

本书是同济大学生命科学与技术学院和图书馆医学信息检索教师多年教学工作的总结、辛勤劳动的结晶,是教师、用户、学生三者通力合作的成果。具体分工如下:陈平老师编写第1章、第2章、第3章第1、2、3节、第8章、第9章。陈善志老师编写第3章第5节、第5章、第7章、第10章和第12章第1节。张轶群老师编写第3章第4节、第4章、第6章、第11章、第12章第2节。同济大学附属同济医院靳令经老师、生命科学与技术学院陈苏、顾华老师结合实际应用对本书部分章节进行了修改。同济大学拔尖人才基地班2011级王玉冰、黄蕾、陈川、景园雅、侯瀚霄、王文、吴悠、王越千等同学根据自己的学习体会对本书编写提供了许多建设性的意见和建议,并对部分章节初稿进行校对。本书出版还得到同济大学研究生院的大力支持,并列为同济大学研究生规划教材(项目编号:2014JCJS031)。科学出版社的同志为本书的出版付出大量劳动,在此一并表示感谢。另外,在编写过程中,本书参考了大量文献资料,由于篇幅所限,不能一一列出,特向有关作者致歉并表示衷心感谢!

本书由多人撰写,文字风格各异,并限于编者水平难免有疏漏和不足之处,敬请专家和读者批评指正,不胜感激。

编 者
2015年1月

目 录

第1章 生物医学信息检索概论	(1)
第1节 信息素养与课程剖析	(1)
第2节 信息检索基础知识	(2)
第3节 计算机信息检索	(10)
第2章 生物医学信息组织	(19)
第1节 信息组织基础知识	(19)
第2节 中国图书馆分类法	(20)
第3节 美国医学主题词表	(22)
第3章 文摘型数据库检索	(33)
第1节 中国生物医学文献数据库	(33)
第2节 PubMed	(44)
第3节 Web of Science	(50)
第4节 Scopus	(60)
第5节 SciFinder	(67)
第4章 中文全文信息检索	(70)
第1节 中文全文期刊数据库	(70)
第2节 中文电子图书数据库	(84)
第3节 中文学位论文数据库	(87)
第5章 外文全文信息检索	(89)
第1节 外文全文电子期刊数据库	(89)
第2节 外文学位论文数据库	(95)
第3节 外文电子图书数据库	(99)
第6章 引文信息检索	(103)
第1节 中文引文信息检索	(103)
第2节 英文引文信息检索	(110)
第7章 学术评价信息检索	(114)
第1节 Journal Citation Reports 数据库	(114)
第2节 Essential Science Indicators 数据库	(115)
第3节 InCites 数据库	(119)
第8章 生物信息学数据检索	(121)
第1节 概况	(121)
第2节 NCBI 提供的生物信息学数据库及其检索	(122)

第3节 常用生物信息学数据及其检索	(139)
第9章 引物设计及优化	(148)
第10章 网络学术信息检索	(160)
第1节 网络搜索引擎	(160)
第2节 开放获取资源	(164)
第11章 循证医学证据检索	(171)
第1节 循证医学基础知识	(171)
第2节 循证医学的证据来源和检索	(177)
第12章 文献管理与信息分析	(188)
第1节 参考文献管理软件	(188)
第2节 文献计量分析和可视化软件	(193)
参考文献	(205)

第1章 生物医学信息检索概论

目前网络发达,信息海量,信息作为继材料、能源之后的第三种重要资源,成为当今社会经济发展的三大支柱之一。在信息的海洋中如何快速获取自己需要的信息资源不仅仅是一个普通“冲浪”者的意愿,更是科技工作者的追求,如何有效地选择和利用信息的技能和技巧变得越来越重要。选择和利用信息的有效科学手段就是信息检索。

在生物医学科学研究方面,学科交叉、渗透、整合越来越明显,生物医学知识和医疗技术更新迅速,生物医学信息数量也以指数级迅速增长。信息有助于医务人员及生物医学科研人员经验的交流、知识的增长和水平的提高。人们对生物医学信息资料的依赖程度进一步加强。如何选择合适的信息源、怎样掌握检索与利用信息资源的基本理论与方法、如何判断信息的真伪、如何评价信息的优劣及学术价值等,已经成为生物医学科技工作者日常生活中的重要内容。

本章主要介绍生物医学信息资源的基本概念、基本知识及文献检索的基本理论,通过本章的学习,学生能够掌握生物医学信息检索基本知识,具备一定信息素养,为以后检索工具的学习奠定基础。

第1节 信息素养与课程剖析

信息素养(information literacy)是伴随着社会信息化的形成和发展而出现的一个名词术语,是指具有检索、分析、评价和利用各种信息源以满足信息需求及制定明智决策的能力,是对个人信息行为能力、独立学习能力及批判性思维能力等的概括性描述。广义的信息素养包括能够判断什么时候需要信息,并且懂得如何去获取信息,评判性地评估和利用信息及信息源等。信息素养一般包括信息意识、信息能力、信息道德等方面。信息意识是信息素养的前提,包括充分认识信息的重要作用,树立终生学习、勇于创新观念,对信息敏感,有洞察力。能迅速有效地确定所需信息的范围,发现并掌握有价值的信息。信息能力是信息素养的重点和核心,是指具有有效地利用各种工具及信息资源获取信息、加工处理信息和创造新信息的能力,包括对所需信息进行高效的存取,确定所需信息范围,批判性地评价、分析信息资源,并将获取的信息与自身的知识加以融合,有效、准确地利用信息资源解决问题。信息道德是指人们在获取、利用信息过程中必须遵循的道德,能理解与信息利用相关的经济、法律及社会因素,并在社会伦理及法律规范之下存取、利用信息。

国内同类图书有多种,内容上既有相似性,又有差异。本书是在多年校内使用的讲义基础上重新整理编写而成,以实用为特色。主要特色有:

- (1) 针对性强:总结多年教学经验和各类学生使用的情况,选择在学习和工作中对生命科学工作者、医学学生及临床医生有切实帮助的数据库和网络资源,列举实例深入浅出地介绍相关内容。
- (2) 时效性强:所选内容既有专业性强的医学传统数据库和网络资源,又结合当前学科研究热点增添了生物信息学、循证医学证据检索、引文分析、文献管理和信息分析等内

容,图文并茂。

(3) 系统性强:从信息检索的基础知识到信息组织、信息检索和获取、信息评价,最后到文献管理,涵盖完整的信息素养标准要求和内容。

总之,本书甄选实用性强、对学习和工作有实际帮助的数据库和网络资源,以最新的内容和界面加以展示。

第2节 信息检索基础知识

一、基本概念

(一) 信息

信息,在科学技术飞速发展的今天已成为人们日常谈论的话题。经济信息、文化信息、政治信息、科技信息;信息社会、信息时代、信息高速公路、信息资源、信息经济等名词充斥各种媒体。那么,究竟什么是信息?作为一般化的信息概念又是什么?至今尚无公认的定义。不同历史时期、不同科学技术发展阶段,各领域学者从不同的角度对此做出不同的解释。

最早把信息作为科学对象来加以研究的是在通信领域。在这里信息被看作有序程度的度量和负熵,是用以减少不确定性的东西。哲学认识论认为信息是物质的普遍属性,是物质世界一切事物运动变化的反映。因为人类社会、自然界、思维乃至宇宙中的一切事物的存在方式及运动状态所形成的固有的规律和特征都由相应的信息表现出来。

随着科学技术和经济的发展,人类认识程度的不断提高,信息的概念也在不断拓展。20世纪80年代前期,美国学者马克卢普指出信息一词在广泛应用中有6种通行的看法:信息是事实、信息是数据、信息是沟通、信息可作为一种能量形式、信息可作为一种商品、信息也可以是知识。

在国内图书馆学界一种观点认为:“信息不是事物的本身,而是由事物发出的消息、情报、指令、数据、信号等当中所包含的内容”。作为日常用语,信息是指消息、情报、指令、密码等,是通过符号(如文字、图像等)、信号(如语言、有某种含义的手势动作、电磁波信号等)等具体形式表现出来的。信息无处不在,无时不有。

从信息产生的根源来看,可分为自然信息和社会信息两大类。自然信息是指自然界中的各种信息和人类生产的物质产品所产生的信息,包括遗传信息、生物信息等。社会信息是指为了特定的目的产生、传递、交流并应用于人类社会实践活动,包括一切由人类创造的语言、符号和其他物质载体表达和记录的数据、消息、经验、知识。

(二) 知识

人类自20世纪末进入知识经济时代后,知识日益成为生产力的主要因素。

1997年世界经济合作与发展组织在《以知识为基础的经济》报告中,依据知识的表现层次对知识进行新的分类。know-what:“知道什么”型知识,关于简单数据与事实;know-how:“知道怎么做”型知识,即关于技能方面;know-why:“知道为什么”型知识,即抽象的、解释层次的;know-who:“知道谁有知识”型知识。每一层知识都是对知识理解层次的扩展和深入。波普尔把知识分为主观知识和客观知识。主观知识是个人头脑中的知识,客观知识则是写、印于载体之上的由陈述构成的知识。科学知识就属于客观知识。储存在图书馆的知

识就是客观知识。

(三) 情报

情报是指被传递的知识或事实,是知识的激活,是运用一定的媒体(载体),越过空间和时间传递给特定用户,解决科研、生产中的具体问题所需要的特定知识和信息。

关于情报的概念,有多种下定义的方法。有学者用拆字的方法,将“情报”两字拆开,解释为“有情有报告”就是“情报”;也有学者从情报搜集的手段来给其下定义,说情报是通过秘密手段搜集来的,关于敌对方外交、军事、政治、经济、科技等信息;还有学者从情报处理的流程来给其下定义,认为情报是被传递、整理、分析后的信息。情报的定义是情报学中一个最基本的概念,它是构建情报学理论体系的基石,是情报学科建设的基础,对情报工作产生直接的影响。情报究竟是什么,时至今日,国内外对情报定义仍然是众说纷纭。据学者统计,如今国内外对情报的定义数以百计,不同的情报观对情报有不同的定义,主要的三种情报观对情报的解释是:军事情报观对情报的解释;信息情报观对情报的解释;知识情报观对情报的解释。除了军事、信息、知识三种主要情报观的情报定义外,还有许多从其他不同的社会功能、不同的角度、不同的层面对情报作出定义的,但在普遍意义上能被多数学者认同接受的情报定义是:情报是为实现主体某种特定目的,有意识地对有关的事实、数据、信息、知识等要素进行劳动加工的产物。目的性、意识性、附属性和劳动加工性是情报最基本的属性,它们相互联系、缺一不可,情报的其他特性则都是这些基本属性的衍生物。

情报具有知识、传递、效用三个基本属性,此外,情报还具有社会性、积累性、与载体的不可分割性以及老化等特性。

(四) 文献

今天我们所说的文献,主要指有历史意义的比较主要的书面材料,广义的文献定义就成了记录知识的一切载体。依据国际定义文献乃是一切情报的载体。

通过一定的方法和手段、运用一定的意义表达和记录体系记录在一定载体的有历史价值和研究价值的知识。所谓文献:文即是文本记载;献就是口头相传的。文献的基本要素是:①有历史价值和研究价值的知识;②一定的载体;③一定的方法和手段;④一定的意义表达和记录体系。人们通常所理解文献是指图书、期刊、典章所记录知识的总和。文献是记录、积累、传播和继承知识的最有效手段,是人类社会活动中获取情报的最基本、最主要的来源,也是交流传播情报的最基本手段。

(五) 检索

检索,即查找、获取。文献信息检索即从大量文献信息集合中查找所需文献信息的活动、过程与方法。按检索对象可分为数据检索、事实检索、文献检索。按检索手段,可分为手工检索和机械检索,后者又可分为脱机检索、联机检索、光盘检索和网络检索。

二、信息资源类型

(一) 信息资源概念

就一般意义而言,资源是指自然界和人类社会生活中的一种可以用以创造物质财富和

精神财富的具有一定的积累和客观存在形态,如土地资源、矿产资源、人力资源等。

信息作为人类发展进程中三个不可缺少的要素之一,同样在无形地创造宝贵的物质财富和精神财富,并以不同的形态存在,因此是人类社会所必需的资源,即信息资源。

信息资源是人类社会经济活动中经过加工处理的有序化并大量积累起来的有用信息的集合。如科技信息、经济信息等。

“所谓信息资源,就是人类社会信息活动中积累起来的以信息为核心的各类信息活动要素(信息技术、设备、设施、信息生产者等)的集合。”(马费成)

(二) 信息资源类型

1. 记录型信息资源 是信息资源存在的基本形式,是信息资源的主体,是记录在各种载体上的具有固定的形式和传播渠道的信息资源总称。包括非数字化信息资源和数字化信息资源。

(1) 非数字化信息资源:包括印刷型信息资源及缩微声像型信息资源。

印刷型是以纸张为载体形成的信息资源。缩微型是以感光材料为载体记录的信息资源,如胶卷、平片等。视听型或声像型是指利用声、光、磁、电等技术记录的信息资源,如录音带、录像带等。

印刷型信息资源数量庞大。按出版形式不同可分为:图书、期刊、科技报告、专利文献、会议文献、国际机构及政府出版物、学位论文、标准文献、技术档案、产品资料等。

图书是对已发表的科研成果、生产技术和工作经验的总结性的论述。具有系统性、概括性、完整性、出版周期长、编著者固定等特点。如教科书、参考书、专著、字典、百科全书、年鉴等。

正式出版的图书,均冠有 ISBN 号,即国际标准书号(International Standard Book Number),是专门为识别图书等文献而设计的国际编号。ISO 于 1972 年颁布了 ISBN 国际标准,并在西柏林普鲁士图书馆设立了实施该标准的管理机构——国际 ISBN 中心。采用 ISBN 编码系统的出版物有图书、小册子、缩微出版物、盲文印刷品等。在联机书目中,ISBN 可以作为一个检索字段,从而为用户增加了一种检索途径。ISBN 由 13 位数字组成,分五个部分:前序列号(978 或 979),组号(国家、语言或区位代码),出版社代码号,书序码和检验码。组号由国际书号中心负责分配,我国大陆的组号为 7。出版社号由国家标准书号中心负责分配,其位数视申请出版社图书出版量多少而异。书序号,由出版社负责管理分配。为了增加数字容量,ISBN 自 2007 年 1 月 1 日由目前 10 位数字扩展到 13 位数字。如 2006 年由中国科学技术出版社出版陈界主编的《医学信息检索与利用》(第三版)的 ISBN 号是 7-5046-1447-5。

期刊是一种定期或不定期的连续性出版物,每期版式基本相同,有固定的刊名,有连续的年、卷、期号。具有更新及时、信息含量大等特点。如:杂志、学报、通报、综述与述评、索引、文摘等。

同图书一样,正式出版的期刊也有国际标准连续性出版物编号 ISSN(International Standard Series Number),是 ISDS(国际连续出版物数据系统)国际中心为在该系统登记的连续出版物分配的号码,以实现对全世界期刊的管理。采用 ISSN 编码系统的出版物有期刊、会议记录等。在联机书目中,ISSN 可以作为一个检索字段,从而为用户增加了一种检索途径。ISSN 由 8 位数字组成,分两个部分:序号和检验码。如中华医学杂志 ISSN:0376-2491。

中华医学杂志(英文版):0366-6999。

印刷型信息资源又可按加工深度不同分为:一次信息资源、二次信息资源、三次信息资源。

一次信息资源即发表在各种载体上的原始论文,是作者对自己工作成果的描述和总结。具有独创性、新颖性。通常包括期刊论文、科技报告、学位论文、专利文献、会议文献、专著等。

二次信息资源是一次信息资源形式上的变化所形成的文献群体,具有条理性、易查找。通常指目录、索引、文摘等,也可称为书目型信息资源。

三次信息资源是对一次信息资源的内容所做的变化而形成的文献群体,具有总结性、提示性。通常包括综述、指南、年鉴、进展等。综述,即 review,是最常见的三次信息资源,是综合、分析、评论某一专题,能全面反映该课题在国内外发展动态的文献。

(2) 数字化信息资源:指来自于计算机存储设备的资源,是网络资源及一切本地的数字化信息资源的统称。数字化信息资源涉及人类生产、生活、娱乐以及其他社会活动的各个方面,是随着人类社会实践的不断发展而不断积累起来的。

2. 实物型信息资源 由实物本身存储和表现的知识信息。如样品、样机、工艺美术品等。

3. 智力型信息资源 主要表现为人脑存储未编码的知识信息,包括人们掌握的诀窍、技能和经验,又称隐性知识。

4. 零次信息资源 指各种渠道中由人口头传播的信息,是人们通过直接交流获得的信息,是信息客体的内容直接作用于人的感觉的结果。具有直接性、纪实性、新颖性、随机性、非存储检索等特征。

三、信息检索技术

(一) 信息检索

根据特定用户在特定时间和条件下的需求,运用某种工具,按照一定的过程、方法和技术,从各种信息资源中查找所需信息,以形成用户多需的信息资源的过程。是一个潜在信息资源显性化的过程。

广义的信息检索包括信息组织和信息检索两个过程。前者是对大量无序信息加工处理使之有序化并形成检索工具和检索文档的过程。后者则是根据需求利用检索工具或检索系统检出信息的过程。

信息检索的实质是将描述特定用户所需信息的提问特征与信息存储的检索标识互相匹配的过程。

信息检索有利于信息资源的充分利用,避免重复劳动,提高工作效率;有利于决策者在把握全局的基础上科学地进行决策,增加决策的有效性和正确性。

信息检索按检索对象可分为数据检索、事实检索、文献检索。

文献检索即通过可检索性检索工具(书目、索引、文摘、全文等)获得间接答案(文献线索或全文),然后进行知识再加工和创新的检索。包括书目检索、全文检索和多媒体检索。

数据检索是从存储有大量数据、图表的检索系统中(手册、年鉴、图谱、药典等)获取数值型信息的检索。检索结果可直接用于知识的分析和再创造。

事实检索是从存储有大量知识信息、事实信息和数据信息的检索系统中(百科全书、名录等)获取某一事物发生、发展进程或过程的检索。如发生事件的时间、地点、人物等。

数据检索、事实检索可通过参考性检索工具直接获得答案(事实或数据)。比如人物信息、统计信息、经济信息、化学、物理数据等。

(二) 信息检索工具

信息检索工具是存贮和检索信息的工具,是将分散在各种信息源上的信息按照一定的编排原则加工、整理编辑形成的二次文献。主要功能是避免信息查找的盲目性、分散性和偶然性,提高检索效率。

依据检索方式可将信息检索工具分为手工检索工具和计算机检索工具。前者指传统印刷型检索工具,后者主要指借助各种计算机查寻的检索工具。按报道方式可分为题录型、索引型、文摘型、全文型检索工具。

题录型检索工具正文提供的是反映文献外部特征的文献内容,即文献的篇名、作者、作者通讯地址、文献来源、文献发表时间和出版单位等。文献内容只能通过文献篇名来判断。

文摘型检索工具正文提供的是以文献内容特征为主要特征的文献形式。文献内容特征即除包括文献外部特征内容外,还包括反映文献内容的水摘。文摘以简明扼要的语句概括文献的主要内容,让读者从中可以了解到文献的大概内容。

全文型检索工具是随着信息技术、通讯技术的发展而产生的。它记录的是文献的全部内容。通过全文检索工具用户即可获得全部信息内容,就像阅读印刷型期刊、图书一样,通常以网络形式存在。

(三) 信息检索工具构成

1. 手工信息检索工具构成 一般的手工检索工具由 5 部分组成:

(1) 编辑说明:即使用指南。包括检索工具介绍、编制目的、收录范围、收录年限、著录规则、检索方法和注意事项等。根据编辑说明,通常能确定检索工具的主题,有利于选择正确的检索工具和检索途径。

(2) 目次表:多数工具书按分类组织编排,是查找正文的依据。

(3) 正文:检索工具的主体。

(4) 辅助索引:包括主题索引、题名索引、著者索引、号码索引等。辅助索引对正文检索起着补充的作用。辅助索引越多,检索途径就越多。

(5) 附录:检索工具内容的补充和参考。通常包括收编的期刊目录、会议名称目录、各种术语、缩写、收藏单位代码等。有的检索工具以单行本出版。

2. 计算机信息检索工具的构成 计算机检索工具一般由检索软件和检索数据两部分组成。不同的检索软件有不同的检索界面,但可以是相同的数据内容。如 MEDLINE CD-ROM 数据库,尽管目前有银盘公司(Silver Platter)、Dialog 公司和 Cambridge 公司三家产品,但它们的数据内容都是一样的,都来自美国国立医学图书馆提供的医学文献。

目前的检索软件通常是在 Windows 平台上操作。一般可以看到菜单栏、功能键、检索输入框、状态栏等。其中菜单栏提供的是基本的操作目标,如打印机设置、显示设置等。功能键的目的是为用户提供快捷的检索方法、打印、存盘、帮助等。状态栏则显示计算机的当前状态。用户若有问题可在“帮助(HELP)”中寻求答案。

(四) 医学常用中外文检索工具

1. 常用中文手工检索工具

(1) 中文科技资料目录:医药卫生:由中国医学科学院医学信息研究所编辑出版。1963年创刊,1978年纳入我国科技检索刊物体系,是《中文科技资料目录》系列刊物的一个分册,是我国目前唯一比较全面报道国内医学文献的印刷型、分类式题录型检索工具。主要收编生物医学类的期刊、学术会议资料、内部资料等。现为月刊,年报道量近6万条。有分类检索和主题检索两种检索方法。

(2) 中文科技资料目录:中草药:是《中文科技资料目录》的一个分册。由国家药品监督管理局主管,天津药物研究院主办,1978年创刊。主要收集国内公开发表和内部发行的中草药文献,以题录形式报道,季刊。

(3) 中国药文学文摘:是国内唯一的药文学文献信息检索刊物。由国家药品监督管理局信息中心编辑出版。1982年创刊。主要收录国内公开发行的700余种相关期刊中的中、西药文献,以文摘形式报道,月刊。

(4) 中国生物学文摘:是中国科学院科学技术文摘系列刊物之一。由中国科学院文献情报中心、中国科学院上海文献情报中心编辑出版。1987年创刊。主要收录国内公开出版的期刊、专著和会议录中有关报道我国生物科学领域的研究成果和进展的文献。以文摘形式报道,月刊。其数据库产品已由中科院上海文献情报中心出版。

(5) 中国医学文摘:是系列性医学文献信息检索刊物。目前有18个分册,分别涵盖医学各学科。主要收录国内公开发行的中央及省、市一级医药卫生期刊上发表的报道各学科的新动向、新进展、新技术和新理论。采用文摘、题录和简介的形式综合报道。

(6) 国外科技资料目录:医药卫生:《国外科技资料目录》是我国出版的查找国外科技文献的系列检索刊物。医药卫生是其中一个分册。1964年创刊。由中国医学科学院医学信息研究所编辑出版。主要收录国外英、日、俄等文种的生物医学期刊中的文献。译成中文,以题录形式报道,月刊。年报道量6万条。

(7) 全国报刊索引·自然科学技术版:《全国报刊索引》是一综合性题录型检索刊物。1955年创刊,1990年起分为哲学社会科学版和自然科学技术版。其中自然科学技术版收录医药卫生文献,月刊。

2. 常用中文文摘型检索工具

(1) 中国生物医学文献数据库(CBM):是中国医学科学院医学信息研究所开发研制的综合性医学文献数据库。1995年问世。该数据库收录自1978年至今的近千种中国期刊,以及汇编、会议论文的文献题录。年增长量约20万条,其中30%文献以文摘形式报道。学科覆盖范围涉及基础医学、临床医学、预防医学、药学、中医学及中药学等生物医学的各个领域。

(2) 中文生物医学期刊文献数据库(CMCC):由解放军医学图书馆研制。收录自1994年以来1200余种中文生物医学期刊的约200余万条文献记录。每年报道量25万条,内容涵盖医药卫生各个领域。30%文献附有文摘。

(3) 中国生物学文献数据库(CBA-CD):由中国科学院上海文献情报中心编辑出版。收录自1986年来至今的国内期刊发表的有关生物学及其相关领域的文献,累计16万余条。每季更新。数据库系汉英双语种文献数据库,每条记录的主题词、关键词、出版物名称、作者姓名、作者单位均有汉英/拼音对照,85%的记录有英文标题,从1995年起60%的记录有

英文文摘,并且英文文摘的比例逐年增加。数据库提供6种检索途径:主题、作者姓名、作者单位、出版物名称、中图法分类以及标题和文摘中英文字(词)的全文检索。主题词索引使用“中国生物学文献数据库系统”编制的《生物学主题词表》标引,关键词索引使用非规范主题词标引,仅作为主题词索引的一种补充。

3. 常用中文全文检索工具

(1) 维普中文科技期刊全文数据库:由重庆维普资讯有限公司制作,是国内最大的综合性文献数据库。数据库包含1989年以来的自然科学、工程技术、农业、医药卫生、经济、教育和图书情报等学科8000余种期刊文献,并以每年100万篇的速度递增。2000年后出版期刊12000余种。有7个专辑:自然科学、工程技术、农业科学、医药卫生、经济管理、教育科学和图书情报。专辑又细分为27个专题。期刊全文采用扫描方式加工,保持了全文原貌。

(2) 中国期刊全文数据库(CJFD)(清华同方数据库系统):由清华同方光盘股份有限公司制作发行,是目前世界上最大的连续动态更新的中国期刊全文数据库。收录自1994年至今国内公开出版的5300种核心期刊与专业特色期刊的全文文献。积累全文文献618万篇,题录1500余万条。分理工A(数理化天地生)、理工B(化学化工能源与材料)、理工C(工业技术)、农业、医药卫生、文史哲、经济政治与法律、教育与社会科学、电子技术与信息科学九大专辑,126个专题文献数据库。数据每日更新。

(3) 万方数据资源系统:包括数字化期刊群、学位论文全文、学术会议全文等产品。由万方数据股份公司制作。

4. 常用英文检索工具

(1) Index Medicus(IM,美国医学索引):是当前世界上最常用的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编辑出版。收录全球范围内68个国家、44种文字、3500种左右的生物医学期刊文献,以题录形式报道。月刊。

(2) MEDLINE CD-ROM:包括IM的光盘数据库。1983年由美国国立医学图书馆创立。收录自1966年以来医学期刊文献。75%文献是英文文献,85%文献含有文摘。数据库包括Index Medicus、Index to Dental Literature、International Nursing Index印刷本全部内容。月刊。

(3) PubMed:由美国生物技术信息中心(NCBI)开发的用于免费检索MEDLINE等数据的网络检索系统。

(4) Elsevier Science公司SDOS全文检索系统:Elsevier Science公司是全球最大的出版商。通过SDOS(Science Direct Onsite)系统为用户提供2000多种电子期刊服务。可在线检索、浏览、打印及下载所需的期刊全文。目前收录的与生物医学有关的期刊共计700余种。

四、信息检索途径、方法和步骤

(一) 检索途径

1. 著者途径(Author Index) 已知某一著者、编者、译者的姓名或机构团体名称,通过著者索引查找文献的途径即为著者途径。

著者索引是指按照文献的著者、编者、译者的姓名或机构团体名称的姓名字顺编制而成的索引,用来查找某个特定的人或团体所写的文献。

利用著者途径查找,简单、快捷,查准率高。但著者的姓名尤其是国外著者的姓名比较

复杂,检索时必须注意。可参考各种文字的译名手册、人名录等工具书。

(1) 姓名次序:欧美人姓名习惯是名在前,姓在后,一个人的姓只有一个,名不止一个。但检索工具中的著者索引和参考文献著录中人名表达方式与习惯相反,姓在前,用全称,名在后,用首字母,姓名之间用空格或逗号分隔。

例:William Henry Harrison(威廉·亨利·哈里森)

↓

William H. Harrison 或 W. H. Harrison(缩写形式)

↓

Harrison, W. H 或 Harrison, WH 或 Harrison-w-h(检索形式)

(2) 团体著者:按原名著录,并加国名以示区别。如美国 FDA(Food & Drug Administration)和中国 FDA。

(3) 姓名前有前缀:姓名前有前缀 de, des, du, la, della, von, van, den, dan, der 等时,将前缀和姓作为一个整体,按字顺排列。词间空格和大写字母不影响排列。如:

van de Donk, H. J

van den Bosch, H

von Werder, K

(4) 族姓:有的著者有家族称呼即族姓,族姓与姓之间无空格,族姓与姓第一个字母均大写。如:

McAda, P. C

MacAdam, K. B

(5) 等级制称号:有等级制称号的著者姓名排在无等级制称号之后,如:

Jons, W. M(琼斯, W. M)

Jons, W. M. II(琼斯, W. M, 二世)

Jons, W. M. Jr(小琼斯, W. M)

(6) 复姓人名:将复姓作为一个整体对待,如:Martin-Smith, M。

(7) 中国人姓名按汉语拼音著录:如:韩启德 → Han Q D。

(8) 同名同姓不同人的情况:鉴于著者索引只将作者名的首字母作为编排依据,因此查找某作者的文献需结合该作者的通讯地址或专业综合考虑。

2. 分类途径(Classified Index) 依据课题在分类体系中的位置,通过分类检索语言查找文献的途径。检索标识即分类号。检索的关键是确定检索课题在分类体系中的确切位置。常用的分类法如《中国图书馆分类法》。

3. 主题词途径(Subject Index) 依据课题的主要概念,确定主题词,通过主题词索引查找文献的途径。检索标识是主题词。

主题词索引是按照主题法检索语言即主题词的字顺编排而成的索引。不同的学科有不同的主题词词表,检索时应注意参考。如《医学主题词表》、《中医药学主题词表》、《中国生物学主题词表》、《汉语主题词表》等。

4. 关键词途径(Keyword Index) 依据课题中具有实际意义并能表达文献概念的词或词组,通过关键词索引查找文献的途径。关键词是未经规范、不受控制但能够表达文献实际意义的词或词组。按照关键词的字顺编排的索引即关键词索引。关键词途径检索的最大优点是无须专业人员指导,读者只要根据课题需要选择关键词即可检索文献。

以上四种检索方法应根据课题已知的条件选择相应的检索途径,确定最简单、最方便、自己最熟悉的途径查找文献。

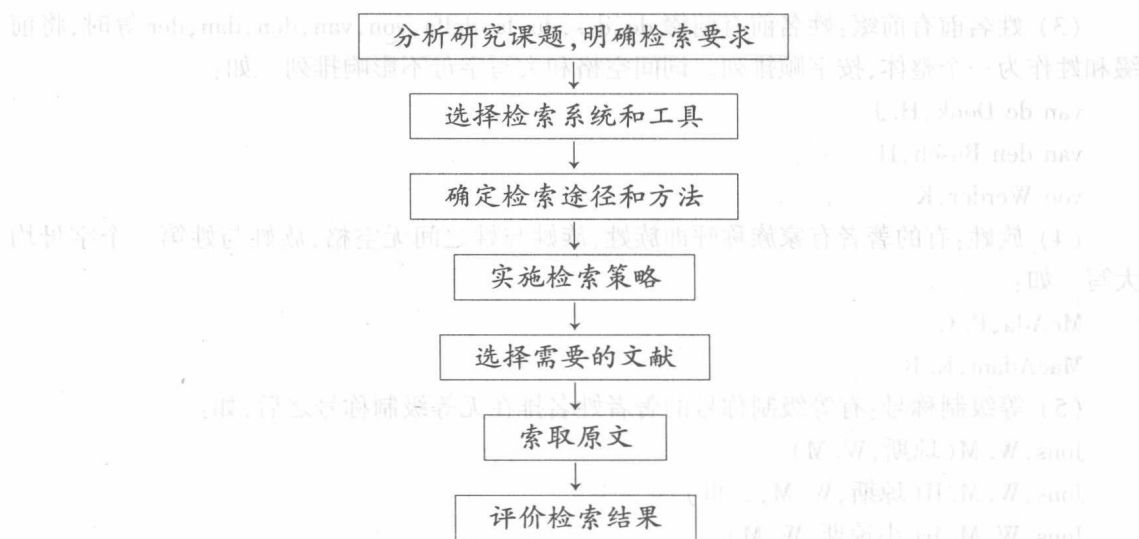
(二) 检索方法

1. 常用法 利用检索工具检索文献的方法。根据课题的需要,可以按照时间先后由远及近或由近及远的查找,也可以在某一确定的具体时间段内查找。

2. 追溯法 利用已知文献后的参考文献追溯查找文献的方法。这是一种比较原始的检索方法,多在无检索工具或检索工具不熟悉的情况下使用。

以上两种方法可根据课题需要交替使用。目的是提高检索效率。

(三) 检索步骤



第3节 计算机信息检索

随着科学技术的不断进步和信息量的急速膨胀,特别是电子技术、网络技术、多媒体技术的发展推动了以信息技术为核心的高新技术的发展,使信息处理、加工、检索从传统模式逐渐向新型模式转变,极大地促进了信息检索技术的发展,从而使计算机信息检索成为现实并得到了飞速发展。

计算机信息检索简称机检,狭义的计算机信息检索是利用计算机查找出符合读者所需信息的过程;广义的计算机信息检索是指读者将检索提问式按照一定的检索指令输入计算机,计算机检索系统将检索提问式与存储在计算机中的信息进行匹配比较,并将符合检索提问式的文献记录输出,由计算机或终端设备显示或打印的过程。

一、计算机信息检索的发展

(一) 计算机信息检索的发展史

自1954年美国海军兵器中心首次将电子计算机技术用于信息检索以来,计算机检索共

经历了五个阶段。

1. 脱机信息检索阶段(20 世纪 50 年代中期到 60 年代中期) 这一时期计算机信息检索系统的工作方式主要是建立一些机读卡片、机读目录,通过计算机对信息进行脱机的批处理检索,虽然只是批处理试验系统,仅包括文献号和少量检索词,但却是计算机技术应用于信息检索的首创。脱机信息检索其数据的存取和数据通信能力都比较差,用户不能与系统进行实时对话,不能及时修改检索策略。因此,人们开始研制更便利的联机检索系统。

2. 联机信息检索阶段(20 世纪 60 年代中期到 70 年代初期) 在 60 年代初人们就开始研制联机检索系统,1965 年美国系统发展公司(SDC)研制出 ORBIT 联机检索软件。1966 年,洛克希德空间与导弹公司建立了 DIALOG 系统,并于 1969 年投入运行。1971 年,美国国立医学图书馆的 MEDLINE(MEDLARS on Line)连接了美国各地的医学机构终端,提供联机检索服务。随着大容量计算机分时系统以及强大功能检索软件相继研制成功,计算机信息检索进入联机检索时代。检索系统由大中型计算机做主机,并通过公共电话线路和许多终端机相连接。用户通过终端可以和主机实现联机对话,检索主机系统所拥有的数据库。在检索过程中可以随时修改检索策略,得到较满意的检索结果,克服了脱机检索阶段中的一些不足。联机信息检索的特点是数据库集中管理,具有完备的联机检索功能,但此阶段因网络技术的限制,数据通信功能较差。

3. 跨国联机信息检索阶段(20 世纪 70 年代中期至今) 此阶段由于卫星通讯、数据通讯网和分组交换通讯网等通讯技术的发展,加之高功能分时计算机、带终端的远程处理系统、廉价的大容量随机存储器等技术的发展,使得联机检索打破了时间、空间的限制,实现了跨国联机检索,成为信息产业的一部分。跨国联机信息检索的特点是检索的范围广、速度快、互动性好。

4. 光盘信息检索(20 世纪 80 年代中期至今) 光盘作为存储设备早在 20 世纪 70 年代就有应用,但作为 CD-ROM(Compact Disk-Read Only Memory)数据库的存储介质,还是从 20 世纪 80 年代开始。这种系统不需要大型计算机,其工作方式主要是使用高密度 CD-ROM 光盘来存储信息,制成光盘数据库,利用配有光盘驱动器的微机安装上相应的检索软件便可实现光盘信息检索。现在,光盘检索还可以利用光盘塔实现多工作站的光盘网络化检索,处于网络中的各个数据库可供多用户同时分享。光盘信息检索系统可提供追溯检索、定题服务、培训用户和作为联机检索前预处理。光盘信息检索不受时间约束,节省了昂贵的联机通讯费,同时因存储量大、价格低廉、使用方便、不依赖于主机和通讯设备,受到了用户的欢迎。

5. 网络信息检索阶段(20 世纪 90 年代至今) 网络信息检索以 Internet 为通信传输通路、采用分布式网络化管理。基于 Internet 的网络化检索系统:它的发展是光纤的利用、多媒体技术的发展、多媒体信息网络、智能计算机和人工智能网络发展的结果。由于 Internet 的迅速发展,在 Internet 上出现了许多新的检索软件和搜索引擎,同时多媒体技术将文字、声音、音乐、图形图像、动画以及视频图像等各种信息加以数字化,利用计算机、通信、信息等多种技术形成多媒体软件,加上先进的触摸频技术,使得计算机信息检索进入了一个多媒体时代,是以住各种检索方式所无法比拟的。基于光盘检索的优点,网络信息检索为了克服网络线路的局限,在 20 世纪末出现了镜像数据库检索,真正做到了在本地检索网络信息资源。

网络信息检索的特点是信息资源内容覆盖面广、形式多样、分布无序、难于结构化和规