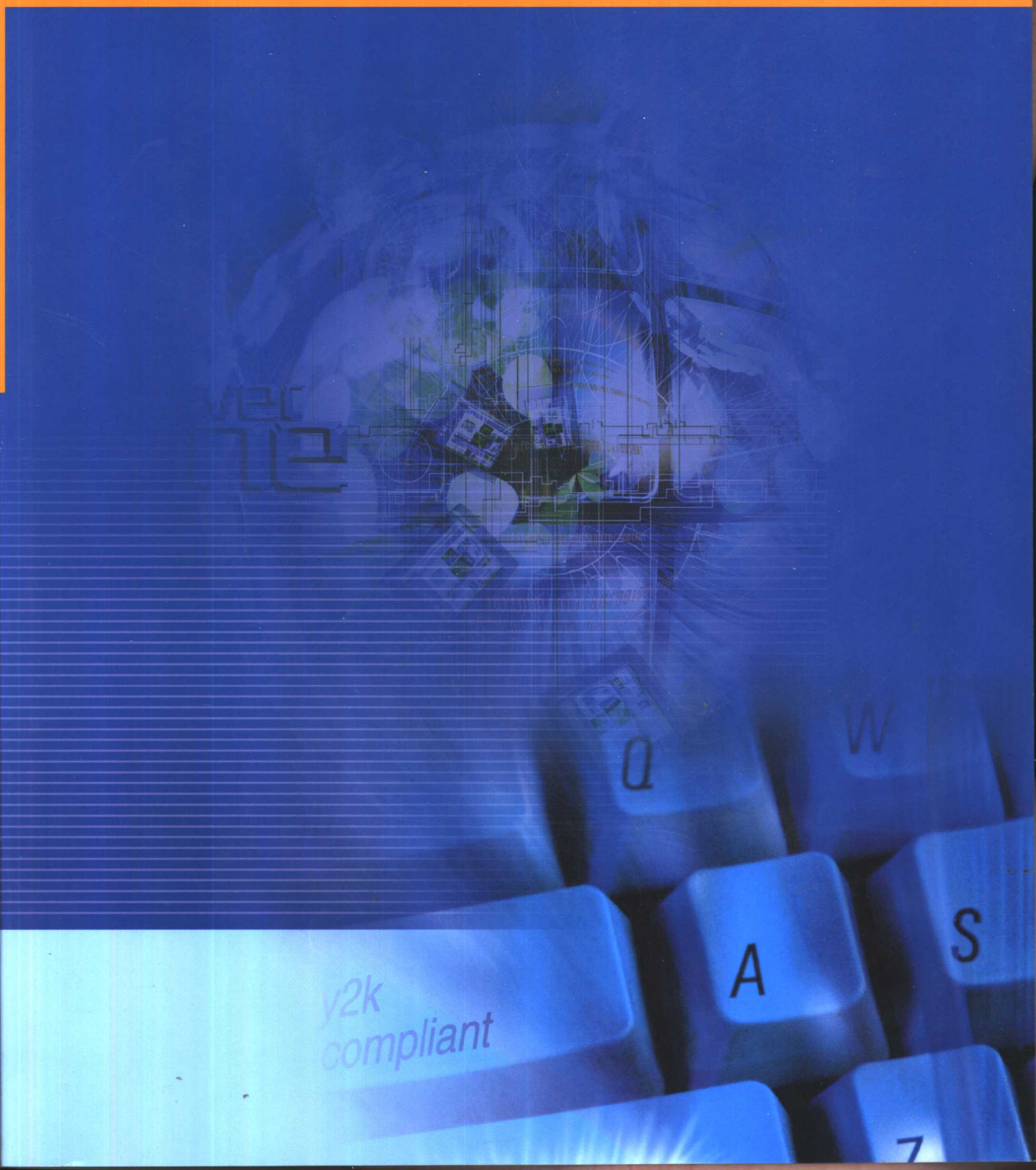


药学信息资源检索

YAOXUE XINXI ZIYUAN JIANSUO

◎孙忠进 何华 主编

东南大学出版社



药学信息资源检索

主 编 孙忠进 何 华

副主编 顾东蕾 钱爱民

编著者 (按姓氏笔画排序)

江 凌 孙忠进 何 华

张晓彤 金 进 顾东蕾

莫莉蓉 钱爱民 章成敏

主 审 蒋光祖

东南大学出版社

内 容 提 要

本书从现代信息技术的实际出发,全面系统地介绍了药学专业文献信息资源检索与利用的理论和方法。在阐述了各种文献信息特点、信息检索原理、参考工具和手检工具知识的基础上,增加了光盘检索、联机检索与网络信息资源的利用等内容。

本书体例新颖、内容丰富、重点突出、特色鲜明、实用性强,既可作为高等医药院校学生教科书,也可作为医药专业技术人员和文献信息工作者参考用书。

图书在版编目(CIP)数据

药学信息资源检索/孙忠进,何华主编. —南京:东南大学出版社,2002.8

ISBN 7-81050-972-1

I. 药... II. ①孙... ②何... III. 药理学—情报检索 IV. G252.7

中国版本图书馆 CIP 数据核字(2002)第 054876 号

东南大学出版社出版发行
(南京四牌楼 2 号 邮编 210096)

出版人:宋增民

江苏省新华书店经销 武进第三印刷厂印刷

开本:787mm×1092mm 1/16 印张:16.75 字数:397 千字

2002 年 8 月第 1 版 2002 年 8 月第 1 次印刷

印数:1-5000 定价:23.00 元

(凡因印装质量问题,可直接向发行科调换。电话:025-3792327)

前 言

我们正面临着飞速发展的时代,而信息已成为这个时代的特征,引人注目。纵观当今全球信息化的大潮,人们再也无法忽视信息的存在及给这个世界带来的巨大变化。信息已成为主导社会经济发展的主要战略性资源,并与物质和能源并列构成现代社会的三大基石。

药学是研究药物的理论与实践的科学,包括药物的发现、发明、加工、生产、供应和使用等诸多环节,是一门综合性的应用科学。药学研究的特点除表现为需要多学科、多专业的科技人员参与,要有先进设备及足够的投入外,也离不开信息的支持。有效地利用信息资源可以为药学研究创造生机和提供保障。

开展信息教育,主动研究信息资源的开发和利用正是迎接这个信息时代挑战的需要。通过药学信息资源检索与利用课程的学习,可以培养学生的信息意识和信息检索技能,使他们在浩如烟海的文献中迅速而准确地获取自己所需要的知识。这对于他们不断汲取新知识,改善知识结构,提高自学能力和研究能力,发挥创新才能都具有重要意义。

近年来,信息技术的发展呈突飞猛进之态势,与信息技术密切关联的计算机检索及国际互联网迅速发展,对信息检索原有的方法或手段都产生了很大的影响。为此,本教材拟在手检和机检相结合方面作些有益的探索,即既不摒弃手工检索的基本内容,又着力加强计算机检索方面的知识,对光盘检索、联机检索与 Internet 网络资源的利用都作了较为充分的介绍,力求做到体例新颖、内容系统、实用性强。

本书由孙忠进、何华任主编,顾东蕾、钱爱民任副主编,参加编写的人员有(按姓氏笔画排列)江凌(第 10 章)、孙忠进(第 1 章)、何华(第 10 章)、张晓彤(第 3 章)、金进(第 2 章)、顾东蕾(第 5、7 章)、莫莉蓉(第 8 章)、钱爱民(第 4、9 章)、章成敏(第 6 章)。蒋光祖任主审。在本书的编写过程中我们始终得到中国药科大学领导的关心和支持,特别是东南大学出版社的有关领导和同志们对本书的编写给予了热情帮助,在此深表谢忱。对本书在编写过程中所参考的文献作者也一并表示感谢!

信息技术的发展日新月异,检索手段的更新速度也越来越快。尽管我们已尽其能力求圆满,但囿于编者水平所限,书中若有错讹或不妥之处,诚请各位专家和读者指正、批评。

编 者
2002 年 7 月

目 录

第一章 信息资源检索概述	(1)
第一节 信息与信息意识	(1)
第二节 信息资源的类型和特点	(3)
第三节 信息检索与检索工具	(8)
第四节 信息检索语言	(11)
第五节 信息检索的方法与步骤	(15)
第二章 药学信息参考工具书	(20)
第一节 工具书基本知识	(20)
第二节 辞典 药典	(21)
第三节 手册 指南	(29)
第四节 百科全书 大全	(33)
第五节 年鉴	(38)
第六节 光谱 图谱	(39)
第七节 丛书	(42)
第三章 中文药学信息资源检索	(44)
第一节 中文药学论文的检索	(44)
第二节 中文药学图书检索	(56)
第四章 外文药学信息资源检索	(60)
第一节 美国《化学文摘》(Chemical Abstracts)	(60)
第二节 美国《生物学文摘》(Biological Abstracts)	(77)
第三节 药学专科检索刊物	(83)
第四节 其他相关检索刊物	(89)
第五章 特种类型的药学信息资源	(96)
第一节 专利文献信息资源	(96)
第二节 其他特种文献信息资源	(112)
第六章 计算机检索基础	(125)
第一节 计算机检索发展简史和现状	(125)
第二节 计算机信息检索系统	(127)
第三节 计算机信息检索技术和策略	(131)
第七章 联机检索	(137)
第一节 联机检索的特点和发展方向	(137)
第二节 国际联机检索的基本方法	(139)

第三节	国际联机检索系统介绍	(140)
第四节	西文联机检索数据库	(150)
第八章	光盘检索	(158)
第一节	光盘检索简介	(158)
第二节	药学中文光盘数据库的检索	(161)
第三节	药学外文光盘数据库检索	(171)
第九章	Internet 信息资源检索	(195)
第一节	Internet 简介	(195)
第二节	网络信息检索工具	(200)
第三节	常用 Internet 数据库检索	(207)
第四节	其他网络医药信息资源的检索与利用	(221)
第十章	药学信息的整理和利用	(229)
第一节	药学信息的鉴别与整理	(229)
第二节	信息的调研流程和方法	(239)
第三节	文献写作	(248)
附 录		(255)
附录 I	BA 主要生物科目分类表	(255)
附录 II	CA 的文种代号缩略词表	(257)
附录 III	黑本式日文罗马字母对照表	(258)
附录 IV	俄文英文字母对照表	(258)
主要参考文献		(259)

第一章 信息资源检索概述

我们生活在一个变幻莫测的信息时代,信息就像我们呼吸的空气一样无所不在。开发和利用信息,已成为我们每个人社会生活的必需。“工欲善其事,必先利其器”,我们要研究现代信息检索,就要掌握其基本功能和原理。明确信息的概念,了解信息资源的类型和特点,以及信息检索工具、检索语言、检索途径,从而便于我们对现代信息检索有一个宏观和一般规律性的认识。

第一节 信息与信息意识

一、信息的概念

(一) 信息

信息一词流行于今,亦见于古。一千多年前,唐代诗人李中在《暮春怀故人》一诗中就有这样的诗句:“梦断美人沉信息,目穿长路倚楼台。”宋朝陈亮在《梅花》一诗里也写道:“欲传春信息,不怕雪埋藏。”上述古诗中“信息”一词,大体上是消息、征兆之意。而信息一词被引入科学领域,则是1948年美国数学家、通信工程师申农(C.E.Shannon)创立信息论开始的。申农在其《通讯的数学理论》一书中给信息下了这样的定义:“信息是二次不定性之差,不定性就是对事物认识不清楚、不知道。信息就是消除人们认识上的不定性。”而另一位信息科学的奠基人,控制论的创始人维纳(N.Wiener)认为:“信息这个名称的内容就是我们对外界进行调节并使我们的调节为外界所了解时而与外界交换来的东西。”由此可见,今天作为事物运动状态和方式的“信息”,其含义与古人所谓“信息”的含义并不完全等同。

信息一词并不神秘,是日常生活中的常见现象的总集,大到宇宙空间,小到微观世界,每时每刻都在不断地发出信息、传递信息。树木的年轮表现了其生长的年龄与发展的信息,海水的涨落表现了月亮围绕地球运行的信息,动物异常状况是地震有可能到来的信息……信息普遍存在于自然界、人类社会和人们的思维之中。

所谓信息是指客观存在着的一切事物,通过一定媒介与形式而向外传播、展示或表现的一种迹象、征兆、信号和消息。简言之,即一切事物(包括物质的或精神的)运动的状态和运动的方式。

当然,本书所指的信息并不是包罗万象的一切信息,而是指人类能够接收和利用的那部分信息,即经过收集、整理、序化了的各类文献信息和社会信息,人们可以通过交流、传递、提问、查询、检索等方式去获取。

(二) 信息的分类

1. 按信息的内容划分

(1)自然信息 自然信息是指宇宙间、自然界客观存在或随机发生的各种事物、事件的信息,是物质运动和生物生存活动的产物,它随着事物的存在而存在,随着事物的变化而变化,是自然界的客观反映。自然信息包括生命信息、动植物“活动”信息及物质物理信息等。人们通过对自然信息的利用可以达到认识世界、改造世界的目的。

(2)社会信息 是指在人类或社会维系生存、生产和发展过程中所产生、传递、利用的信息,是人和人之间发生交往的产物,也是人类从事一切活动的基础。社会信息包括经济信息、政治信息、法律信息、文化信息、军事信息、科技信息、社会生活信息等。

2. 按信息的来源划分

(1)直接信息 事物的存在状态和运动形式本身在人们头脑中的直接反映是直接信息,这种信息是人的眼、鼻、嘴、手、耳、皮肤等感觉器官直接接触客观事物,进而对客观事物作出反映而生成的信息。

(2)间接信息 关于事物存在状态与运动形式的描述(如经过加工整理后的数据、资料、理论、观点等)则是间接信息,是可以通过历史资料查阅,各种信息传播媒介传递,以及对未来事物发展趋势预测等所获取的信息。

3. 按信息的产生过程划分

(1)原始信息 原始信息是事物第一次发出且未经人们作任何处理的信息,如科学发现、科学发明中首次揭示事物的状态或运行规律的信息及科学试验、经济活动中的原始数据。原始信息本身不受人的意识干扰,它是事物状态或方式的客观反映。

(2)再生信息 再生信息是以原始信息为基础,经过加工处理以后而产生的信息,也称之为加工性信息,如文摘、索引、分析报告等。有的再生信息是对原始信息的简化描述或浓缩所得,有的则是对原始信息进行重新组合或分析计算而成。

二、信息意识与信息需求

面对信息社会,人们能否自觉地掌握信息检索及利用已有的知识与技能去猎取有关的信息,其关键在于有没有强烈的信息意识。何谓信息意识?信息意识是指人们从思想上对于信息在经济、社会、事业发展中的地位、价值、功能与作用的认知,是指人们对信息重要性的认识程度和需求信息的迫切程度,以及捕捉信息、分析信息、判断信息和吸收信息的自觉程度。换言之,信息意识就是在充分认识信息价值的基础上,对信息具有特殊敏感的一种主体意识。

(一) 信息意识与信息环境

在当今的信息社会,人们面临着两种境况:其一,日新月异的信息技术作为一种专门的知识与技能要求人们不断学习与利用它,以适应其发展;其二,信息产品渗透到各个领域,信息产出量呈指数级增长,知识更新迅速倍增,这就是当前所处的信息环境,人们一般称之为“信息爆炸”。信息量的急剧增长还引起现代社会的“信息危机”,一方面是社会信息的迅速累积、高速增长,新知识、新成果不断涌现;另一方面,个人、组织、团体乃至整个社会的信息传播、处理与吸收能力发展较慢,使大量有价值的信息不能发挥有效作用。人们为了在竞争中立于不败之地,谁都想要先掌握有价值的信息,唯恐自己所知道的东西越来越少和越来越陈旧。现代社会的信息环境与竞争状况已经给人们造成压力,使人们普

遍感到信息匮乏,因而产生危机感。对此所采用的对策就是培养自己高度的信息意识及信息敏感性,时时处处从信息的角度去观察周围的一切,及时注意捕捉各类信息,能从分散、无序的文献信息中发现价值高、内容新的信息。

(二) 信息意识与人才素质

在信息社会,信息意识强的人都是把搜集信息作为事业的生存之本,取胜之道。这无疑是对人的基本素质提出了新的要求。美国劳动部提交的一份关于 2000 年人才需求的报告中提出,信息时代,人们特别需要五个方面的能力,其中使用信息的能力占有重要的地位。也就是说,具有信息意识和获取信息能力将成为人们在未来社会里生存和发展的基本素质之一。在信息社会,人们若缺乏获取信息的意识和能力,就会成为“信息盲”,甚至丧失一些基本生活能力。而信息意识的觉醒和获取信息技能的提高,能有效地改善文化素养和思维方式,增强创新能力,及时调整知识结构,有效适应信息社会的要求。

(三) 信息意识与药学创新

药品是知识密集型的高科技产品。药品的研制、开发不是一项简单的工程,是创新知识的过程。现代药学研究是在传统的医药学基础上不断地吸收科学技术的新信息、新技术而发展起来的。20 世纪中叶以后,分子生物学、生物工程技术、电子计算机技术等方面的发展,使药物设计、作用机理研究和药品生产进入了一个新的历史发展阶段,新药的研究和开发向新领域拓展。新中国成立 50 多年来,新药研究基本上走的是仿制之路,自主创新药物较少。随着世界经济越来越趋向地区化、全球化,国际、国内市场的竞争越来越激烈。没有新产品,只靠仿制,将会连自己国内的市场都可能守不住。因此,我国必须尽快走创制新药的道路。

由此可见,我国的药品研制、开发和生产目前面临着严峻的形势和挑战。创制新药是一项复杂工程,需要多学科、多专业科技人员参与,还要有先进设备及足够的投入等条件,其中信息是不可忽略的重要环节之一。我们要有敏锐的信息意识,及时了解和掌握国内外在药品研制方面的有关信息,包括药品研制方面的历史状况、现有动态、最新成果、发展趋势和前景,以及可供开发的环境和可利用的条件等,为发展我国的医药事业,为创制有自主知识产权的新药,在方方面面创造必要的条件。

第二节 信息资源的类型和特点

一、信息资源

所谓资源指的是一切可被人类开发和利用的物质、能量和信息的总称。它广泛地存在于自然界和人类社会中。信息最初并不被认为是资源,随着信息在社会与经济发展中的重要性与日递增,信息资源化才逐渐被认识,信息最终成为人类经济社会发展的一种重要的可供利用的资源。所谓信息资源则是指经过人类开发与组织的信息的集合。

信息资源同其他资源一样是多种多样的。社会信息资源一般有三种存在形式:一是存贮在人脑记忆中,人们通过语言交谈、讨论等形式进行交流传播;二是存在于文献中,用

文字、图形、符号等手段记录在一定的物质载体上;三是存在于实物中,人们通过实地观察、采集样品等形式进行传播交流。这三种形态分别称之为口头信息资源、文献信息资源和实物信息资源,它们构成了社会的信息资源总体。在这三者中,文献是信息资源的主要组成部分。当然在信息技术飞跃发展的今天,文献资源越来越多地表现为数字化,并且具有海量存贮、网络化传递以及多途径检索或智能检索的功能,具有世界范围内共享的特征。电子信息资源现已占有相当重要的地位。

二、信息资源的类型

为了更好地开发和利用信息资源,应首先了解信息资源的类型。我们以文献资源为主体从收录的内容、加工程度、获取的难易程度、载体形式来进行简单的分类。

(一) 以信息资源的内容形式划分

1. 图书

图书是一种传统的、比较定型的信息媒体,数量庞大、流通广泛、影响深远持久,是用以表达思想、保存知识、积累经验、传递信息的工具。其特点是反映的知识内容比较完整系统和成熟,论述全面。不过,传统印刷型图书出版周期长、体积大、更新慢,电子版图书的出现将弥补这一缺陷。

2. 期刊

期刊又称杂志,是指那些定期或不定期连续出版,有固定的名称和版式,使用连续的卷、期号或年、月顺序序号作为时序标识的一种连续出版物。

期刊按其内容性质或使用对象可以分为学术类期刊、检索类期刊、政论类期刊、文艺类期刊等。由于期刊具有出版周期短、报道时差短、能及时反映当前科技水平等特点,是人们传递信息、交流学术思想所使用的最基本、最广泛的载体。据统计,从期刊方面获取的科技信息约占整个信息来源的 65%。

目前,在网络资源中新兴的网络刊物发展很快,倍受人们关注。

3. 专利文献

专利文献包括专利申请书、专利说明书、专利公报和专利检索工具,以及与专利有关的一切资料。

4. 标准文献

标准文献是指按规定程序制定,经权威的管理机构批准的一整套须在特定范围内执行的规格、规则、技术要求等规范性文献。

5. 科技报告

科技报告也称研究报告或技术报告,它是指有关某项科学研究的正式报告或研究进展情况的阶段性总结和实际记录。

6. 会议文献

会议文献指的是在国内外各种会议上所产生的正式和非正式的资料汇总,是围绕会议的中心议题,在会前提提供的发言预印本或发言预摘、会上发表或散发的论文,以及会后整理出来的有关资料的总称。

7. 政府出版物

政府出版物是指各国政府部门及其所属机构发表或出版的各种文献总称,一般可分为行政性文件和科技文献两大类。

8. 学位论文

学位论文通常是指高等学校、科研单位攻读学位人员为获取学位而撰写的学术论文,分为硕士论文和博士论文。

9. 产品资料

产品资料是各国厂商为了推销产品而印发的企业出版物或免费赠送的商业宣传品,通常在一定程度上包含着有关产品技术信息。

10. 科技档案

科技档案是指单位在技术活动中所形成的技术文件、图纸、图片、原始技术记录等资料。包括任务书、协议书、技术指标、审批文件、研究计划及方案、调研材料等。

(二) 以信息资源的加工程度划分

1. 一次信息资源

一次信息资源是指报道新发明、新创造、新知识、新技术的原始信息,是人们研究或创造性活动成果的直接记录。如期刊论文、研究报告、专利说明书等,它是加工二、三次文献信息的基础。

2. 二次信息资源

二次信息资源是对一次信息资源进行加工整理后产生的一类文献信息。如书目、题录、简介、文摘等检索工具,它是查找一次信息资源的桥梁。

3. 三次信息资源

三次信息资源是在一、二次文献信息的基础上,经过分析、综合而编写出来的文献信息。如综述、述评、文献指南等。

(三) 以信息资源获取的难易程度划分

1. 白色信息资源

白色是一个明亮的色调,因而人们将公开发行,较易获得的文献信息形象地称为白色文献信息资源,如正式出版的图书、期刊等。

2. 黑色信息资源

黑色与白色相对,指看不见的事物状态,因而人们将处于保密状态,极难获取的文献信息称为黑色文献信息资源,如未解密的政府文件、技术档案等。

3. 灰色信息资源

灰色介于黑色、白色之间,因而人们把既不易找到、又非不可能得到的文献喻为灰色文献信息资源,如会议资料、未公开的学位论文等。

(四) 以信息资源的载体形式划分

1. 纸质型信息资源

纸质型信息媒体是以手写和印刷技术为手段,以纸张记录信息的载体形式。包括铅印、油印、胶印、木版印刷、激光照排印刷等形式的出版物,也包括手工书写以及复印机复制、传真传递、计算机打印等文献,通常是以纸质材料为介质。它有悠久的历史,至今仍广

为流传,成为传播文献信息的主要形式。其优点是读取方便,不受时空的局限;其不足是存贮信息密度低,占据空间大,笨重,管理比较困难。

2. 缩微型信息资源

缩微型信息媒体是以感光材料为存贮介质,以缩微照像为记录手段产生的一种信息媒体形式,它是原始信息媒体的缩微复制品。它包括缩微胶卷、缩微平片等。缩微型媒体的优点是信息储量大、体积小、价格便宜、保存期长,一张 105mm × 148mm 的缩微平片,缩小 1/24、1/42 和 1/48 时,可分别存入 98、325 和 420 页资料。超缩微片一般每张可存 3 200 页,而特超缩微片每张可存 22 500 页。其缺点是不能直接阅读,阅读时必须借助于阅读设备。

3. 视听型信息资源

以磁性材料或感光材料为存贮介质,借助特殊的机械设备,直接记录声音和图像,并通过视听设备存贮和播放信息知识的信息媒体形式,如唱片、录音带、幻灯片等。它依感官接受功能可划分为视觉资料、听觉资料和声像资料。其优点是直观性强、形象逼真,可直接表现那些难以用文字描述的事物。其缺点是必须借助一定的设备才能使用。

4. 电子型信息资源

它是指采用电子手段并以电子形式存在,利用计算机及现代通讯方式提供信息的一种新兴信息媒体,原称机读型信息资源。它通过计算机对电子格式的信息进行存取和处理。即采用高技术手段,将信息存储在磁盘、磁带或光盘等一些媒体中,形成多种类型的电子出版物,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。它具有信息存储密度高、存取速度快的特点,并且具有电子加工、出版和传递功能,其出版物形式分为单行版电子出版物、联机数据库型电子出版物和网络型电子出版物。如各种光盘型数据库、联机数据库、电子期刊、电子图书、电子报纸、电子会议、电子邮件、电子公告等。还可以进一步分为文本型、多媒体型、超文本型。

文本型主要是以数字形式记载文字信息,让用户通过计算机来阅读,文本型的媒体也可进一步分为题录型、文摘型和全文型,题录型通常只提供文献的著者、题名、原文的地址;而文摘型的数据库在提供题录的基础上,还可检索到文献的摘要。随着信息技术的发展,全文型文献计算机检索阅读应运而生。

多媒体型是能够把文字、图形、图像、声音动画和视频图像采集于一体,通过计算机进行综合处理和控制的媒体,它改变了信息的表示方法,对同一信息能同时采用图、文、声等多种形式表示,是有多多样性与直观性和人机交互的友好界面。

超文本型(包括超媒体型)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,自由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方式。

三、药学信息资源的特点

(一) 载体多样、传递快捷

现代信息技术对信息资源载体形式及传递方式的影响是最直接、最根本的。在 20 世纪 60 年代初,计算机只是作为信息处理和文献编排工具应用于生产印刷型文献。随着现代信息技术的不断发展,印刷型文献信息一统天下的局面被打破,出现了各种形式的电子

出版物,增加了磁盘、光盘等载体形式和电话、传真、网络等传递方式。由此而形成信息资源的多类型、多媒体、跨时间、跨行业、跨地域的快速增加。很多信息资源以印刷型、光盘型和网络型三种形式同时出版发行,这种记录知识载体的形式多样化,使其性能更具有科学性。目前,以印刷型信息资源、磁带型信息资源、光盘型信息资源、网络型信息资源为主且这四种信息资源并存的格局已经形成,现代信息技术的发展,加速了信息转换,丰富了信息资源载体形式和传递方式。

(二) 内容丰富、数量激增

现代信息技术对信息资源内容的变化也有不可低估的影响。现代科学技术和社会经济的发展为信息资源内容的拓展和数量的增多提供了动力和条件。由于现代信息技术的应用,信息生产周期缩短,信息资源数量迅速增多,属传统信息资源的图书、期刊、特种文献等正式出版物,全世界正以每分钟 2 000 印张的速度在增长,使人们应接不暇。在浩如烟海的信息中,医药文献信息占有很大比重,约占 1/4,目前世界上生物医药刊物已达 22 000 多种,每年发表的医药论文近 300 万篇。与此同时,光盘型信息资源和网络型信息资源的数量日益扩大并持续动态地增长着,网络上既有学术性专业论文的正式信息交流,也有非正式信息交流,如电子邮件、电子布告牌、网络会议中所包含的信息内容以及各种学术团体、企事业单位、政府部门、行业协会等单位通过正式出版物系统无法得到的灰色信息。

(三) 分布广泛、交叉分散

在印刷型文献信息一统天下的时代,资源的社会分布相对集中,图书馆、信息所、档案馆是信息资源的主要分布点,信息服务部门的信息主要来自于出版社、编辑部、书店和图书发行公司。如今信息资源的社会分布异常分散,数量众多的信息资源广泛地分布在各类社会机构、社会组织以及大部分家庭中。信息服务部门的信息除了来源于出版社、编辑部、书店外,还可以来自于计算机硬件和软件公司、数据库开发公司,甚至是遍布世界各地无法确定的千千万万服务器上。而个人可利用网络传播自己的研究成果和其他信息,也成为信息来源之一。现代科学技术综合与交叉的特点,使某一学科专业内容和相关信息分布范围极为广泛,药学除与医学、化学关系密切外还涉及生物学、生物化学、生理学、环境科学、农业、工业以及管理科学、市场营销等学科,专业信息的分散和交叉重复,无疑增加了开发和利用信息资源的难度。

(四) 历史悠久、蕴义精深

我国药学研究至今已有几千年的历史,在漫长的发展过程中,积累了大量的文献信息,由于这些信息高度的经验性和实用性,得到人们的普遍重视和尊崇。在当今“人类回归大自然”的趋势下,人们越来越重视从天然药物中研究和开发新药。根据 WHO 资料,未来制药业创制的来自中药的新药可能由目前的不到 5%,快速增长到 29% 以上,一些欧美大医药公司纷纷成立天然药物部,不惜重金搜集中医药信息。祖国古代药学研究成果主要保存在历代遗留下来的中医药文献典籍之中,现存本草专著约 278 种,还有些中药信息散见于综合性中医著作中。由于年代久远,有些书籍散佚,加之时代变迁、文字古奥,词义艰深,这都为获取中药信息造成了比较大的困难,因此积极开发和利用中药信息也是当

今重要的任务之一。

第三节 信息检索与检索工具

一、信息检索的原理

(一) 信息检索的概念

一般认为,信息检索是根据特定的需要,利用检索工具或其他技术手段,从大量的信息中,迅速准确地查找出所需信息或单元知识的一种方法或过程。信息检索有广义和狭义之分,广义的信息检索包括信息的存贮与检索,是指将信息按一定的方式组织和存贮起来,并根据人们的需要找到有关信息的过程和技术。狭义的信息检索仅指检索,即从信息的集合中找到所需信息的过程。

(二) 信息检索的基本原理

信息检索的基本原理是指研究信息存贮和检索的基本原理和方法。信息存贮是将信息存贮到检索工具中去,组成检索系统,提供检索手段。信息检索则是根据需要,从检索工具或检索系统中将信息查找出来的逆过程。

信息存贮的过程一般包括:搜集原始信息进行筛选;对所选出的信息进行主题分析,得出主题,将主题转换成主题词或分类号(即标引);著录其信息的内容特征和外部特征;按照一定的排检方法,编成检索工具,这就是信息的输入过程。信息检索的过程一般包括:对检索课题进行主题分析;确定检索标识;根据标识符到检索工具中去查找,即将检索系统中的信息标识与检索提问的检索标识进行匹配,当二者达到一致时,即为“命中”信息而予以提取或输出。

(三) 信息检索的类型

根据不同的标准,可以将检索分为不同的类型。

1. 根据检索内容分类

(1)文献检索 是指从一个文献集合中查找出特定课题的文献信息的行为、方法或程序。

(2)事实检索 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。

(3)数据检索 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统计数字、图表、化学结构式、计算公式等。

三种检索虽然各有特点,但检索的基本原理和方法没有本质区别,在实际工作中也不能截然分开,所提供的信息常常是有关联的和互补的,但比较而言,文献检索最为基本,其应用也最为广泛。

2. 根据检索手段分类

(1)手工检索 是指人们以手工方法,利用活页式、卡片式、期刊式、附录式等检索工具来处理和查找信息的行为。

(2)机器检索 是指人们借助机械、光电、计算机、网络等技术设备处理和查找信息的行为。主要包括穿孔卡片检索;缩微检索;使用磁带、磁盘的脱机、联机检索;光盘检索以及利用互联网的网络检索。

手工检索灵活性高、费用低,在今后一段时间仍是重要的检索手段之一。但计算机检索、网络检索发展很快,以其特有优势越来越受到人们的重视,因此,现代科技人员不但要掌握传统的手工检索、机械检索技术,更应掌握计算机检索技术和网络检索技术。

二、信息检索工具的形式

由于信息的类型多种多样,人们对信息检索的角度、广度和深度也不一致,因此产生了多种检索工具。

(一) 印刷式检索工具

1. 期刊式检索工具

以期刊形式刊行,有统一的年、卷、期号,它报道信息快而新,主要用于查检最近发表的文献信息,也可用来回溯查检,这种检索工具适用范围广,利用率也很高。

2. 单卷式检索工具

它以某一专题内容编制,并累积多年与之有关的文献信息。这种检索工具收录的文献信息按学科分类集中,报道量大,内容系统全面,时间和地域范围广,适合用来回溯查检专业性强的文献信息。

3. 卡片式检索工具

以特定规格的卡片记录文献信息有关款目,再以一定的方法把卡片排列起来,即成为卡片式检索工具。它编排灵活,可随时累积组合、剔旧更新,用户使用也颇为方便。

4. 附录式检索工具

它不是一个独立存在的出版物,一般附在图书、期刊或论文之后,以辅助索引、引用书目、参考文献等形式出现。

(二) 缩微式检索工具

是一种摄制在缩微胶卷或胶片上通过缩微阅读机才能使用的检索工具。另外还有一种 COM(Computer Output Microfilm,计算机输出缩微胶片)式检索工具也属于缩微式检索工具。

(三) 机读式检索工具

将检索工具的信息以数字方式存储在磁带、磁盘和光盘等介质上,通过电子计算机输出设备和电信网在视频终端上显示出来,它具有贮存容量大、存取速度快等特点。

三、信息检索工具的类型

(一) 传统型检索工具

所谓传统型检索工具主要是指以纸质材料为载体,以手工检索为特征的一些工具书。这类工具书按功用可分为两大类。一类是资料型的,为人们提供的是各种实际工作中需要的具体资料,如专业术语、字词的解释,它本身就包含读者所需要的信息。这一类出版

物有辞典、药典、百科全书、手册、大全等,我们将在专门章节详细介绍。另一类是检索型的,是从书刊文献信息外部特征和内容特征出发,用特定的编排形式和检索方法编制而成的专供人们迅速查找信息线索的出版物,如目录、索引、文摘、综述等。

1. 目录

目录是以文献的出版单元为著录对象,系统揭示其名称、著者、出版者、出版时间以及收藏者、内容提要的检索工具。目录揭示了文献的基本特征,提供了图书本身的有关情况,通过目录可以掌握文献信息发展的基本状况、了解某文献在该学科发展中的地位和作用,其作用在于揭示藏书、指导阅读。

目录有很多种类,按职能划分有国家书目、出版目录、馆藏目录、联合目录等;按文献种类划分有图书目录、报刊目录等;按物质形态划分有卡片式目录、书本目录、机读目录等。

2. 索引

索引是将若干文献中具有检索意义的特征,如书名、刊名、篇名、人名、地名、主题、语词等项目分别摘录出来,按一定方式加以编排并逐一注明出处,以供查检的一种检索工具。

索引从文献的外部特征和内部特征分析、选择、标引文献所含的信息内容,并按一定顺序编排,可满足人们多种的检索要求,从而提高了检索深度和检索效率。人们通过索引能掌握学术动态,开阔视野,启发思路。因此索引的使用比较广泛,其他检索工具也大多附录各种辅助索引。

索引的种类很多,根据其功能,主要分为篇名索引和内容索引两大类。

3. 文摘

文摘是将文献的重要内容进行简略而确切的摘述,并按一定著录规则和排列方式系统编制而成的检索工具。

文摘以简明扼要的文字陈述文献的主要论点、数据、结论等,并注明出处,是原始文献的浓缩。文摘条目的著录方式和编排方式类似于索引,但每一条目下附有该文献的内容提要。所以,文摘除了具有索引的某些性能外,还有其独特的功用:扼要报道最新的理论动态、研究成果和其他信息,节省读者的时间和精力,参考价值比较高。

文摘的种类较多,根据文摘摘录方式及其作用的不同,可以划分为题录式文摘、指示性文摘、报道性文摘。

(二) 电子型检索工具

电子型检索工具是一种依托现代计算机技术和网络技术,以磁带、磁盘、光盘等为存储介质,以数字化的形式通过电信号、光信号传输信息的检索工具。我们可以从数据库形式和网络形式来进一步了解。

1. 数据库

数据库是以某一特定方式编制和存贮数据资料的一种新型检索工具,是电子信息检索的基础。按国际上通用的分类方法,数据库分为以下三大类:

(1)参考数据库 指能指引用户到另一信息源获取原文或其他细节的数据库,包括书目数据库和指南数据库。

(2)源数据库 指能直接提供所需原始资料或具体数据的数据库,包括数值数据库、

全文数据库、术语数据库和图像数据库。

(3)混合数据库 指能同时存储多种类型数据的数据库。

2. 计算机信息网络

计算机信息网络是指通过远程通信方式进行计算机信息交换与数据库的存取而形成的一种系统。这也就是网络信息检索工具,可以分为三种:

(1)分类目录型 分类目录型检索工具提供按类别编排的因特网站目录,其检索方法为分类目录浏览检索。

(2)搜索引擎型 搜索引擎型检索工具提供关键词查询网站及网页信息,其检索方法为关键词查询检索。

(3)混合型 混合型检索工具兼有搜索引擎和分类目录两种检索方式,既可直接输入关键词查找特定信息,又可浏览分类目录了解某个领域范围的信息。

电子型检索工具一般数据完备、检索手段先进,受到广泛的欢迎。目前许多传统型的检索工具正在逐步向电子型转换,发展十分迅速。

第四节 信息检索语言

一、检索语言的概念和种类

(一) 概念

语言是人际交流的工具,信息检索语言与自然语言一样,都具有表达思想,相互交流的功能。所不同的是,前者仅限于人与检索工具之间,或人与计算机数据库之间的信息交流和沟通。具体地说,信息各方面的特征都可以用一种特定的语言(各种词、词组、短语或符号)记录下来。人们以这些词或符号作为标识,把相同的标识集中排列起来,使之成为一个供组织和检索信息使用的标识体系。检索语言与自然语言之不同还在于在概念的表达上具有惟一性、排他性,消除了自然语言中存在的多义词、同义词等不适于检索的缺点。总之,检索语言是信息检索时人们用来描述信息特征和检索提问的一种专用语言,了解检索语言是熟练运用信息检索工具的基础。

(二) 种类

检索语言有多种分类方法,主要有:

1. 按其学科专业划分,有综合性语言和专业性语言。
 2. 按表达信息的特征划分,可分为表述信息外表特征的检索语言和表述信息内容特征的检索语言。
 3. 按组配方式划分,可分为先组式检索语言和后组式检索语言。
 4. 按结构原理划分,可分为分类检索语言、主题检索语言、代码检索语言。
- 目前使用较普遍的是分类语言中的体系分类法和主题语言中的叙词法、关键词法等。

二、分类检索语言

分类是人们以事物的属性为标准对事物进行区分的一种常用方法。分类检索语言是