

NONGYE XINXI JIAN SUO

信息检索丛书

农业 信息检索

◎ 包 平 主 编

东南大学出版社

信息检索丛书

农业信息检索

主 编	包 平		
副 主 编	郑江平	吕 娟	吴贤奇
编 者	(按姓氏笔画排序)		
	毛莉菊	包 平	乔 颖
	刘景会	吕 娟	孙文海
	吴贤奇	张洪艳	张 彬
	陈国秀	屈卫群	郑江平

东南大学出版社
• 南京 •

内 容 提 要

本书概要阐述了信息检索的相关概念,基本原理与技术;重点介绍了国内外农学、生物学及相关领域数据库、电子图书和电子期刊,详细叙述了各检索工具的使用方法和技巧;突出了在网络环境下数字资源的使用和数字图书馆的概念与发展。

本教材内容新颖、重点突出,适合本科生、研究生作为教材,同时也可作为教师及科研人员检索文献的参考工具书。

图书在版编目(CIP)数据

农业信息检索 / 包平主编. —南京:东南大学出版社,
2003. 8

ISBN 7—81089—254—1

I. 农... II. 包... III. 农业—情报检索
IV. G252. 7

中国版本图书馆 CIP 数据核字(2003)第 072309 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 溧阳市晨明印刷有限公司印刷

开本:700 mm×1000 mm 1/16 印张:22.25 字数:474 千字

2003 年 8 月第 1 版 2003 年 8 月第 1 次印刷

印数:1~5000 册 定价:29.00 元

(凡因印装质量问题,可直接向发行科调换。电话:025—3795801)

前 言

当海量信息的存储和组织变得日益容易,当多媒体等不同类型的文献的快速检索成为现实,当网络的带宽变得不再拥挤,未来信息存储和检索技术将可以使人们在任何时间、任何地点获取各类信息成为可能。在信息化社会,人们获取信息的技巧和能力,决定着其综合竞争能力。

获取信息的能力来源于对信息组织和检索基本知识的掌握,以及大量的检索实践。十年前,信息检索教材大多是手工检索工具的编排和使用,今天,信息检索教材更多是传授基于网络的计算机检索原理。在众多农业信息检索教程中,本书具有如下特点:首先,将目前国内正在使用的农业与生物学及内容相关的数据库、电子期刊、电子图书等内容进行了介绍,图文并茂,举例确切,文字组织简洁易懂,充分体现时代和学科的要求和特色;其次,首次将“INTERNET 与数字图书馆”独立成章,目的在于引导读者接受一种全新的理念和方式,这种信息的数字化、网络化、全球化正不以人们意志为转移地改变着我们的学习、生活的方式;第三,以读者检索目的的不同,列举了以了解科研动态、科技查新、撰写综述、查找专利信息等多个完整的典型案例,目的在于帮助一些自学者和初学者在实际检索时参考应用。

本书简化了信息的存储、组织、检索等方面的技术和原理,强调了实用性,主要适宜于非图书情报专业的本科生和研究生使用,同时兼顾继续教育的需要,供在职人员自学或作为检索工具书。

本书的编写得到南京农业大学图书馆、信息科技学院,浙江大学图书馆,扬州大学图书馆,华南农业大学图书馆,东北农业大学图书馆,山西农业大学图书馆,河北农业大学图书馆,天津农学院图书馆,南京农业专科学校图书馆,江苏省科技情报研究所等单位的支持,东南大学出版社的张慧老师投入大量的精力,南京农业大学图书馆的何建新、郑萍,信息科技学院的汤亚芬等同志校订了部分内容,在此一一表示深深的谢意!

由于时间仓促,加之编者水平的限制,不足之处请读者与同仁不吝赐教。

包 平

2003年6月于南京卫岗

目 录

1 信息检索基础	(1)
1.1 信息检索的概念范畴	(1)
1.1.1 信息的概念	(1)
1.1.2 文献的概念	(2)
1.1.3 知识的概念	(3)
1.1.4 信息、文献、知识的关系	(3)
1.1.5 信息检索的概念	(4)
1.2 信息检索的目的、意义	(4)
1.3 信息检索的类型及特点	(6)
1.3.1 按存储和检索的对象分类	(6)
1.3.2 按信息存储和检索的方式和技术分类	(6)
1.3.3 按信息组织的方式分类	(7)
1.4 信息检索的发展趋势	(8)
2 信息检索原理与技术	(10)
2.1 信息组织及检索原理	(10)
2.1.1 信息检索语言	(10)
2.1.3 常用信息检索语言	(11)
2.2 信息检索技术	(15)
2.2.1 联机检索	(15)
2.2.2 光盘检索	(16)
2.2.3 网络信息检索	(17)
2.2.4 全文检索	(18)
2.2.5 多媒体检索	(18)
2.2.6 超媒体及超文本检索	(19)
2.3 信息检索的方法与步骤	(20)
2.3.1 信息检索的方法	(20)
2.3.2 信息检索的途径	(21)
2.3.3 信息检索的步骤	(23)
2.4 信息检索效果	(26)
2.4.1 信息检索效果	(26)
2.4.2 提高查全率和查准率的方法	(28)

3 国内农业文献检索	(30)
3.1 印刷型检索工具.....	(30)
3.1.1 印刷型检索工具概论.....	(30)
3.1.2 常用印刷型检索工具.....	(32)
3.2 农业及生物学数据库.....	(39)
3.2.1 中国农林科技文献数据库.....	(39)
3.2.2 中国生物学文献光盘数据库.....	(39)
3.2.3 中国生物医学文献数据库.....	(43)
3.3 综合及相关数据库.....	(47)
3.3.1 中国期刊网.....	(47)
3.3.2 维普《中文科技期刊数据库》.....	(58)
3.3.3 万方数据.....	(64)
3.3.4 电子图书.....	(84)
3.3.5 国务院发展研究中心信息网	(100)
4 国外农业文献检索	(106)
4.1 印刷型检索工具	(106)
4.1.1 美国《化学文摘》(Chemical Abstracts,CA)	(106)
4.1.2 美国《生物学文摘》(Biological Abstract,BA)	(111)
4.1.3 《BA/RRM》(生物学文摘/报告、评论、会议)	(112)
4.1.4 国际农业和生物科学中心文摘(CABI)	(112)
4.1.5 《食品科学与技术文摘》(Food Science and Technology Abstracts,FSTA)	(114)
4.1.6 美国《农业文献题录》	(115)
4.1.7 联合国粮农组织《农业索引》	(116)
4.1.8 美国《环境文摘》(Environment Abstracts,EA)	(117)
4.1.9 美国《EI》	(119)
4.2 农业及生物学数据库	(121)
4.2.1 国际三大农业综合数据库	(121)
4.2.2 生物学及食品科学数据库	(142)
4.3 综合及相关数据库	(160)
4.3.1 化学与工程数据库	(160)
4.3.2 电子期刊	(173)
4.3.3 其他数据库	(202)
4.4 文献传递服务	(222)
4.4.1 文献传递服务的概述	(223)
4.4.2 我国的文献传递	(223)

4.4.3 国外文献传递服务机构	(226)
5 特种文献检索	(230)
5.1 会议文献	(230)
5.1.1 会议文献概述	(230)
5.1.2 会议文献的类型	(230)
5.1.3 会议文献的检索	(231)
5.2 科技报告	(234)
5.2.1 科技报告概述	(234)
5.2.2 科技报告的收藏与管理	(235)
5.2.3 科技报告的检索	(235)
5.3 学位论文	(243)
5.3.1 我国学位论文的检索	(243)
5.3.2 国外学位论文的检索	(244)
5.4 专利文献	(250)
5.4.1 专利文献概述	(250)
5.4.2 中国专利文献及其检索	(252)
5.4.3 国外专利文献及其检索	(256)
5.5 标准文献	(260)
5.5.1 标准文献概述	(260)
5.5.2 标准文献的检索	(263)
6 Internet 与数字图书馆	(271)
6.1 Internet 信息检索	(271)
6.1.1 Internet 概述	(271)
6.1.2 Internet 搜索引擎	(277)
6.1.3 Internet 免费农业信息资源	(288)
6.2 数字图书馆	(289)
6.2.1 数字图书馆概述	(289)
6.2.2 数字图书馆定义	(290)
6.2.3 数字图书馆的产生背景与发展现状	(292)
6.2.4 数字图书馆技术体系构成及相关技术	(293)
6.2.5 数字图书馆雏形	(300)
7 参考工具书	(311)
7.1 参考工具书及其使用的基本知识	(311)
7.1.1 参考工具书与工具书、普通图书	(311)
7.1.2 参考工具书的一般结构与排检方法	(312)
7.1.3 参考工具书的质量鉴别	(314)

7.2 参考工具书的类型及特点	(315)
7.2.1 字典、词典	(315)
7.2.2 类书、政书	(316)
7.2.3 百科全书	(317)
7.2.4 年鉴	(319)
7.2.5 手册	(320)
7.2.6 名录	(321)
7.2.7 图录、表谱	(323)
7.3 重要参考工具书介绍	(323)
7.3.1 印刷版参考工具书举要	(323)
7.3.2 电子版参考工具书举要	(326)
8 信息检索实例	(328)
8.1 以了解科研动态为目的的检索	(328)
8.2 以撰写综述为目的的检索	(335)
8.3 以查新论证为目的的检索	(338)
8.4 以查找专利文献为目的的检索	(342)
参考文献	(346)

1.1 信息检索的概念范畴

自古以来,人们随时都在自觉不自觉地接收、传递、存储和利用着各种信息。在人类社会的演变和发展过程中,人类的信息活动从来没有间断过,信息一直在积极地发挥着人类已经意识或还没有意识到的重要作用。但是,信息作为一个科学概念以及科学研究对象,却只有四五十年的历史。20世纪以来,随着科学技术的空前进步,信息和材料、能源一起构成了现代社会文明的三大支柱,使人们对信息的认识和研究进一步深刻和拓展。然而,到目前为止,人们从不同的学科领域角度、不同的认识层面,对信息的概念作了许多不同的解释,至今还没有统一、完备的定义。

1.1.1 信息的概念

信息可以说是一种十分广泛的概念,它在自然界、人类社会以及人类思维活动中普遍存在。信息,是人类在认识和改造客观世界中所依赖的和使用的一种既非物质又非能量的东西,它的表现形式可以是消息、信号、数据、情报或知识等;同时,信息又是客观世界中一切物质的属性,它反映的是物质之间的普遍联系。既然信息具有普遍存在的属性,同时又是物质存在和运动的反映。因此,本书从对信息检索领域及其应用的角度出发,给信息一个比较广泛的、容易接受和理解的定义:信息是一切事物存在方式和运动状态的客观反映。

要把握、理解信息的概念,我们应了解信息的几个共同特征:

(1) 客观性

信息是客观存在的,它源于客观存在的物质及其运动两大特性。信息的客观性还表现为它是以物质的客观存在为前提的,即使是主观信息,如决策、判断、指令、计划等,也有其客观实际背景,并以客观信息为“原料”,受客观实践的检验。

(2) 依附性

信息总是依附于一定的物质载体而存在,需要某种物质承担者。如果不依附各种适当的载体,信息的含义和价值则不能传递和发挥。声音、语言、文字、颜色、图像、各种符号、光电磁、生物等各种信息信号、纸张、胶片、磁带、光盘、人的大脑等等,无一不是信息的载体。

(3) 可传递性

信息的产生是同信息的传递联系在一起的,信息在传递过程中发挥作用。信息的传递过程是一个重复使用的流通过程,在这一过程中,信息的占有者不会因传递信息而失掉信息,一般说来,也不会因多次使用而改变信息的自身价值。

(4) 可塑性

信息可以加工处理,可以压缩、扩充和叠加,也可以变换形态。在流通和使用过程中,经过综合、分析、再加工,原始信息可以变成二次信息和三次信息;原有的信息价值也可以实现增值;为了有效地交流和传递,借助于先进的信息技术,文本、图像、数字、语言等各种形态的信息均可实现互相转换。

(5) 时效性

现代社会中,信息的使用周期迅速缩短,信息的价值实现取决于及时地把握和运用信息。信息是活跃的,不断变化的,及时地获取有效的信息将获得信息的最佳价值,如时效性很强的天气预报、经济信息、交易信息、科学信息等。不能及时地使用最新信息,信息的价值就会随其滞后使用的时差而贬值。

(6) 共享性

信息的共享性主要表现在同一内容的信息可以在同一时间由两个或两个以上的使用者使用,而信息的提供者并不失去所提供的信息内容和信息量,它是信息资源的广泛提供与利用的基础。

1.1.2 文献的概念

在我国古代,“文献”包括历代的历史文件和贤者两个方面。后来,“文献”一词的内容发生变化,一般泛指具有史料价值的文章和图书。目前,“文献”一词广泛应用于图书馆学、情报学、目录学等研究领域。我国国家标准《文献著录总则》中则用简洁的语言规定:“文献:记录有知识的一切载体。”从定义中可以看出:文献既不是知识本身,也不是记录知识的物质载体本身,而是两者的共同组合;知识是文献的内容,载体是文献的外在形式,文献正是知识内容与载体形式的统一体。

随着科技的发展,随着记载知识的新载体和手段的不断出现,存储和表达人们思想的物质载体不断更新、变化,从竹简、羊皮、石板、丝帛、纸张发展到用感光介质、磁性介质及其他新型载体;另外,文字也不再是表达思想的惟一手段,图形、声频、视频等同样是表达思想、传递感情的重要手段。因此,可以说文献的类型非常繁多,五花八门,而且还不断有新的文献形态出现。在信息检索领域,可按一定的划分标准将文献划分成不同的类型。如,按文献载体物理类型可将文献分为:印刷型文献、缩微型文献、机读型文献、视听型文献等;按文献载体的出版类型可将文献分为:图书、报纸、期刊、特种文献等,其中特种文献包括科技报告、会议文献、专利文献、学位论文、标准文献、政府出版物等等;按文献的加工层次可将文献分为:零次文献、一次文献、二次

文献、三次文献等等。因此,我们不能把文献的内涵仅仅局限于常见的纸质形式的信息载体上。

值得一提的是,虽然将文献定义为“记录有知识的一切载体”,但随着新的文献载体不断出现,“一切载体”比较笼统,有的载体是固态的、可见的,有的载体是动态的、不可见的。如声波、光波、电磁波,未来的信息高速公路,数据经压缩后能以比现在高数千倍的速度在光纤上快速传递。因此,有人提出“信息通道”的概念,它可以分为存储型和传播型。文献应属于存储型的固态的载体,而不是同样可承载和传递信息的图文电话、语音信箱、电子邮件、网络等。所以应该可以把文献理解为“记录有人类精神信息的、且便于存储或传递的人工固态附载物”。这里强调的是那些在载体形式上便于保存和传递的纸质和非纸质的固态载体,而非瞬时信息的附载物或其他实物,这两点限定可以帮助我们把文献置于实用的、可操作的范围内。否则,过分扩大文献概念的外延只能给文献信息工作带来困难,使理论和实际脱离。

1.1.3 知识的概念

知识是人类社会实践经验的总结,是人的主观世界对客观世界的概括和如实的反映。知识是人类通过信息对自然界、人类社会以及思维方式与运动规律的认识,是人的大脑通过思维重新组合的系统化的信息的集合。因此,人类不仅要通过信息感知、认识和改造世界,而且要根据所获得的信息组成知识。根据当代著名科学哲学家卡尔·波普尔的“世界3”理论,知识可以分成两种,即主观知识和客观知识。主观知识主要是指人们头脑中的经验、观点、思想,在文字产生以前,知识主要以主观知识的形式出现,但随着信息载体发展,主观知识实现主体外存储,从而产生了通过文字、图像等记录于载体的客观知识,它是人类的一个重大进步。主观知识可以随着人的死亡而消失,而客观知识可以代代继承,从而实现跨越时空的交流、传递与开发利用,因而客观知识的作用和意义不可低估。人类文明的进步,主要是客观科学知识的作用,它也是当代社会信息资源的主体结构。

1.1.4 信息、文献、知识的关系

通过以上对信息、文献、知识概念的定义和解释,关于信息、知识和文献之间的关系我们可以归纳为:信息是广泛存在于自然界、人类社会、人类思维之中的一切事物的存在方式和运动状态的客观反映;知识是人类社会中经过人类智能加工的系统化的信息;文献则是记录人类信息、知识、情报的一切载体。同时,信息、知识、文献又可以不停地互相转换的。

关于信息、文献和知识之间的范畴关系如图1-1所示。

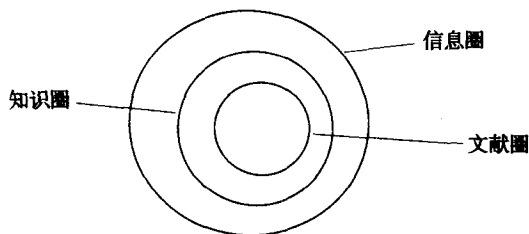


图 1-1 信息、文献、知识的关系

因此,只有科学地了解和认识信息、文献、知识三者的概念和关系,才能更好地把握信息,才能使信息更好地为科学技术、经济和社会发展服务。

1.1.5 信息检索的概念

信息检索是信息工作的一项重要内容,是沟通广大信息用户和各种信息源的主要渠道,也是大量信息进行广泛传递的集中表现形式,最能体现信息存储的目的。信息检索一般是指查找满足特定需要的过程。虽然这种查找行为很早就萌芽于图书馆的参考工作之中,但信息检索的概念直到 20 世纪 50 年代以后才产生。这一概念的产生是由于:① 现代记录下来的知识急剧增长而需要给以合理的存储;② 人们查寻的着眼点更多地指向包含在文献中的信息而不是文献本身;③ 电子计算机应用到信息工作领域使人们拥有强有力的信息处理手段。

(1) 广义的信息检索

从广义上理解,信息检索是指将信息按一定方式组织和储存起来,并根据信息用户的需要找出有关信息的过程。所以它通常被称为“信息存储与检索”,包括各种文献信息的加工存储和检索利用两方面的内容,由两个方向相反而又相互依存的工作过程构成。广义的信息检索概念是对信息工作者而言的。

(2) 狭义的信息检索

狭义的信息检索是指广义的信息检索的后半个过程,即从信息集合中找出所需要信息的过程,相当于人们所说的信息查检。在实际工作中,也往往把日常的信息加工和存储视为内部准备工作,而狭义地把信息的查检视为信息检索。狭义的信息检索的概念通常是针对信息用户而言的,即根据信息用户的特定要求查找所需信息的全过程。

1.2 信息检索的目的、意义

信息检索在信息用户与信息源之间起媒介作用,它是联系信息生产者与信息需求者的中间环节,是信息交流和传递的重要过程,是提高文献信息利用率和科研效率

的重要手段,也是合理利用人类智力资源的重要措施。概括起来,信息检索的意义主要体现在以下几个方面:

(1) 信息检索是获取知识的有效途径

在当今文献和知识急剧增长的信息时代,如何在浩如烟海的文献信息流中迅速找到所需文献信息,并加以合理有效的利用,是摆在科研工作者面前的一道重要课题。据测算,人类知识总量在 19 世纪每 50 年增加一倍,20 世纪初每 10 年增加一倍,20 世纪 70 年代每 5 年增加一倍,80 年代几乎是每 3 年增加一倍。而且,文献信息的分布极不平衡,同一学科的论文分散在几十种,甚至几百种期刊上,导致文献的查找越来越困难。因此,如何以最少的精力、最短的时间充分占有文献信息,成了人们亟待解决的实际问题,而信息检索正是有效解决这一问题的最好途径,它可以帮助人们快、准、全地获取所需信息,最大限度地节省查找时间,使文献信息得以充分的利用。

(2) 信息检索能够避免重复研究和走弯路

科研的任务是为了创造新的知识成果,它要求对某一课题或某一领域的认识及判断应是前所未有的。因此,任何科研项目都必须建立在完全的、充分的信息检索基础之上,否则很容易造成重复研究,导致人力和物力的严重浪费。据统计,我国低水平的重复研究现象比较严重,尤其是省级及以下科研项目,重复率达到 50%。可见,要进行有价值的科学研究,必须全面地获取有关文献信息,及时了解各学科领域出现的新问题、新观点,这只能依赖信息检索才能实现。

(3) 信息检索可节省劳动时间

信息检索是研究工作的基础和必要环节,成功的信息检索可以起到事半功倍的效果;尤其是随着信息技术的更新发展,研究者可以在很短时间里完成信息调研,从而节省劳动时间,提高工作效率。

(4) 信息检索能协助管理决策

科学的决策,源于对信息资料的充分占有。管理决策必须依赖信息咨询才能保证其科学性、公正性;信息咨询成功的基础则是通过科学合理的信息检索获取大量有用的信息。因此,信息检索是国家、企业、单位和个人等获得信息的重要途径。

(5) 信息检索能提供科学的方法

有人曾用“信息爆炸”一词来描述当今社会中信息生产与信息载体数量的迅猛增长,在浩如烟海的信息资料之中盲目地找寻自己所需要的信息,自然是一件非常困难的事情。信息检索则为人们提供了一套较完整的开发、利用信息资料的方法。包括信息检索工具的选择,信息检索策略的制定,信息检索手段的选择等。

(6) 信息检索是科研工作的组成部分

科学研究必须以继承前人成果和借鉴他人工作为基础。一般包括资料准备阶段和研究阶段。其中,资料准备阶段要占用科研人员全部科学研究的 30%~60%的时间,上下幅度随着信息检索效率的高低发生变化。高效的信息检索不仅能使科研人

员减少查找资料的时间,而且还能够通过信息检索获得启发,对研究工作具有极大的帮助作用,最终缩短科学研究的时间。

1.3 信息检索的类型及特点

信息检索按不同的标准可划分为不同的类型,不同类型有各自不同的特点。

1.3.1 按存储和检索的对象分类

(1) 文献检索

文献检索是指以文献为查找对象的检索。用户为了获取某个课题、某个问题、某个事物的相关信息,而去查取记载这些相关信息的文献,就是我们所说的文献检索。它查找的对象是文献,只有当相关文献被找出后,才能得到相关的信息。但这些相关的信息并不能回答用户怎样去解决这个问题,只能对用户解决问题起参考作用。因此,文献检索是一种相关性检索,一种非确定性检索,它的检索原理与方法一般比较复杂,需经过专门培训才能掌握。

(2) 数据检索

数据检索是指用户利用某些检索工具去查找某个确定性的数值型数据而进行的检索。它的目的是为了查找某些确定性的数据,而且这种数据是数值型的。如铁的密度,某个建筑的高度,某个单位的确切人数等。它所利用的检索工具是参考工具书,如手册、年鉴等,或者是相关的数据库。根据其概念我们知道它是一种确定性的检索,能为用户提供直接的答案、确定的结果。因此,数据检索是一种确定性的检索,检索结果要么正确、要么错误,不存在是否能提供多少参考价值。

(3) 事实检索

事实检索是指用户为了确定某个事件或事实发生的时间、地点和过程等方面的信息而进行的检索。它的检索结果可视为一种非数值型数据。因此,有人也将它归入数据检索类中。事实检索直接解决用户的问题,结果要么正确、要么错误,它也是一种确定性检索。

1.3.2 按信息存储和检索的方式和技术分类

(1) 传统信息检索

传统信息检索一般称为手工检索。手工检索是指人们通过手工的方式来存储和检索信息。检索过程是完全由人来进行的行为,专业信息管理人员用人工的方式来完成信息的存储,用户通过手翻、眼看、大脑判断来完成检索过程。根据人们的手工检索实践,又可将它分为专业性检索和非专业性检索。非专业性检索是指人们直接

针对原始文献进行检索,不是利用检索工具来进行。这种检索是盲目的,不仅费时费力,而且检索结果不理想,查全率和查准率都很低。现在许多没有经过检索培训的人还是常用这种非专业的检索方式。专业性检索包括追溯法、工具法和综合法。追溯法是指人们在查出部分与课题相关的文献后,再通过这些文献的参考文献,找到参考文献的参考文献,由此可追溯查出更多的与课题有关的文献。工具法是指人们利用检索工具进行信息检索,这是一种较专业的方法,检索结果往往比较理想,但用户须经过培训才可熟练掌握。综合法则是指追溯法与工具法的综合利用,人们可以先通过工具法找到部分文献,再利用追溯法查找更多的有参考价值的文献,两种方法反复交叉使用,直到检索结果满意为止。

(2) 现代信息检索

由于现代人们已广泛利用计算机进行信息存储和检索工作,因此现代信息检索一般就是指计算机检索。计算机检索是指人们利用数据库、计算机软件技术、计算机网络以及通信系统来进行信息的存储和检索。计算机的产生使信息检索发生了革命性的变化,人们利用计算机大大提高了信息的处理和检索能力,并能在短时间内完成大量的信息检索工作,而且检索效果很好,使人们在信息服务方面得到了巨大的改善。计算机检索的原理比较简单,但有很强的检索技巧性。因此,它也是一种专业性的检索,用户需进行专门的培训方能较好地完成任务。

计算机检索按其检索方式又可分为光盘检索、联机检索和网络化信息检索。光盘检索再分为单机光盘检索、光盘联机检索。计算机的出现促进了信息检索革命性的发展。计算机检索从过去的单机检索到现在的网络化检索,使人们能在短时间内获取大量的、有价值的信息。现代信息检索技术已成为每个人必须掌握的技能之一。

1.3.3 按信息组织的方式分类

(1) 目录检索

目录检索是指以目录的形式来组织信息,形成目录型检索工具,然后用户利用目录来进行检索。目录通常有卡片式目录、书本式目录、机读目录、联机公共检索目录。通过目录只能检索到文献的外部特征信息如书名、著者、出版者、出版地、页码、分类号、文献收藏单位等书目数据,使人们了解到文献的各种特征,然后根据这些外部特征再去查找文献,从而得到原文。

(2) 题录检索

题录检索是指以题录的形式来组织信息,形成题录型检索工具,它类似于目录检索,但其检索结果不是文献的外部特征信息,而是文献中某篇文章的外部特征信息。

(3) 文摘检索

文摘检索是以文摘的形式来组织信息,形成文摘型检索工具,它是在题录的基础上增加了篇章的内容摘要。这样人们不仅能查到文献中某篇文章的外部特征信息,

更能大概了解篇章的内容是什么。

(4) 全文检索

在计算机检索中,由于计算机存储和处理信息的能力很强,因此可以将文献的全文存储在数据库中,并且可以针对全文进行检索。全文检索可以直接用检索词在全文中进行匹配,如果全文中包含相关的检索词,则认为这篇文章满足用户的检索要求,是用户所需的文章。

(5) 超文本检索

传统的文本是线性的,用户需按顺序阅览。超文本则是非线性的,是一种网状结构。超文本的基本组成元素是节点和节点间的逻辑链接链,节点存储信息,节点中的信息和逻辑链接链被联系在一起,构成相互交叉的信息网络。当用户激活某个节点时,计算机就会在屏幕上显示出相应节点的信息。与传统文本的线性关系不同,超文本检索强调中心节点之间的语义链接结构,依靠系统提供的复杂工具作图示穿行和节点展示,提供浏览式查询。

(6) 超媒体检索

超媒体是在超文本的基础上发展而来的,早期的超文本以文字为主,随着多媒体技术的发展,开始将图像、图形、视频、动画、音频等超媒体信息载体用来建立超链接。因此超媒体的存储对象超出了文本范畴信息的存储结构,已经从单维发展到多维,存储空间亦在不断地扩大。信息的浏览方式多采用多窗口形式,一个窗口显示一个节点的内容。

1.4 信息检索的发展趋势

传统的手工方式处理信息的能力是有限的,现代信息社会中信息每年成几何倍数增长,人们对信息服务的要求也越来越高,不仅要求信息传递迅速,达到即时传递,同时对信息新颖性和广泛性的要求也越来越强,还要得到全球范围内不同语种的最新的、最准确的信息。手工检索方式处理信息的速度和能力已远远不能满足人们对信息的需求。因此,依托通讯网络的现代计算机信息检索成为信息检索发展的不可逆转的趋势。现代计算机信息检索是以计算机为核心的人们基本的实践活动之一。现代信息检索在近几年来发展非常迅速,由起初的单机检索系统到现在的网络化检索系统,从单一文字检索到多媒体检索等,人们处理信息和开展信息服务的方式和能力正在日新月异地发生着深刻的变化。

现代的计算机检索系统不仅能处理文字信息,更能处理大量的图像、图形、音频、视频和动画等信息,通过检索系统的语言转换,使人们可以得到不同自然语言环境的信息。而且信息能在第一时间从信源到达人们的手中,达到即时传递。

计算机信息检索的智能化交互功能越来越强。实现多媒体信息检索以来,利用软件技术,人们可以同检索系统进行自然语言的人机交流。当人们将自己检索内容

以自然语言的方式提供给计算机检索系统后,检索系统就会根据相应的要求进行检索,并将结果输出。在一些发达的国家,计算机检索系统已经实现了多媒体 3D 动画式的较强的智能化交互功能,用户只需通过语音系统将提问直接读给计算机,计算机就能根据用户的语音提问进行检索。同时,通过虚拟的咨询人员与用户进行交互式回答,直到用户得到满意的检索结果为止。这种多媒体智能化的服务方式将是现代信息检索系统发展的主要方向之一。

在实现智能化交互式检索功能的同时,检索系统也在不断地简化用户界面,以最友好、最简便的操作让用户进行检索。检索界面越来越大众化,不仅经过培训的人能够使用,而且尽可能让儿童或老人都能够使用。就如现代许多开发人员所提倡的“只要会点鼠标的人就会使用我们的系统”。因此,友好的检索界面也是一个检索系统的主要发展方向之一。

在过去,检索系统一般是面对专业的检索人员。但在现代社会,信息检索已成为人们日常生活和工作的一件“日常事”,检索系统的对象也由过去的专业检索人员逐渐变为最终用户。检索对象的改变,决定了检索系统在对信息进行存储管理和检索的原理和技术上的变化。上述的检索界面越来越友好,操作越来越简单就是一个例证。因此,检索系统在今后的发展上,面对的应是越来越多的最终的信息用户,信息检索大众化的方向已不可逆转。

随着网络化检索系统在各国的发展,越来越多的国家认识到实现信息共享和交流的重要性。在互联网日益发达的今天,人们对计算机检索系统必须具有跨国家和地区的信息检索功能要求也越来越强烈。许多人都希望能在一个简单的、统一的检索平台下,以最少的操作、最短的时间检索到尽可能多的国家和地区的信息。因此,检索系统在保证界面友好、操作简便的同时,提供跨地区的信息服务功能也越来越重要。在已进入信息化社会的美国,OCLC(Online Computer Library Center)就是一个全球化检索系统的雏形,它不仅将美国许多信息服务单位的众多的检索系统联系在一起,而且发展到欧洲和亚洲等国家和地区,形成了辐射全球的服务网络。

在现代社会,为了满足人们对信息的需求,计算机信息检索正朝着多媒体、交互式、智能化、大众化、网络化、全球化方向发展。

(孙文海,吴贤奇)