

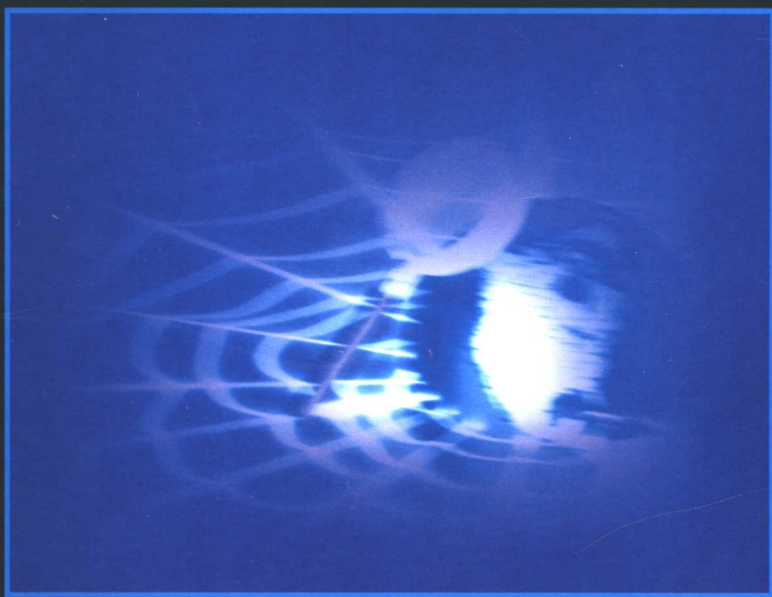


高等院校公共基础课系列教材

浙江省高等教育重点建设教材

数字资源检索教程

章云兰 万跃华 舒炎祥 编著



科学出版社

www.sciencep.com

高等院校公共基础课系列教材
浙江省高等教育重点建设教材

数字资源检索教程

章云兰 万跃华 舒炎祥 编著

科学出版社
北 京

内 容 简 介

本书系统地介绍了数字资源及其检索的基本知识以及中文文献数据库、外文文摘数据库、外文全文数据库、事实数值型数据库以及电子期刊、电子图书、电子报纸、特种文献等数字信息资源的检索方法与技巧,并就网络学术资源的收集方法进行了介绍。此外,还对信息资源的收集技巧和分析利用、信息管理工具 E-Reference Tool 以及论文写作与投稿指南等作了具体介绍。本书基本反映了国内外主要的数字资源,内容新颖实用、全面系统,编排结构合理,显示了最新的检索方法。

本书可作为高等院校本科生、研究生的信息检索课教材,也可作为科技工作者、教师、图书情报人员的参考用书。

图书在版编目(CIP)数据

数字资源检索教程/章云兰,万跃华,舒炎祥编著. —北京:科学出版社, 2005

(高等院校公共基础课系列教材)

ISBN 7-03-016123-8

I. 数… II. ①章… ②万… ③舒… III. 数字技术-应用-情报检索-教材 IV. G354.4

中国版本图书馆 CIP 数据核字(2005)第 090728 号

责任编辑:陈晓萍 / 责任校对:耿 耘

责任印制:吕春珉 / 封面制作:飞天创意

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>

双青印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2006年2月第 一 版 开本: 787×1092 1/16

2006年2月第一次印刷 印张: 23 1/2

印数: 1—7 000 字数: 557 000

定价: 30.00 元

(如有印装质量问题,我社负责调换<双青>)

销售部电话 010-62136131 编辑部电话 010-62138978-8003

前 言

随着 Internet 的迅速普及与发展,人类社会的信息化进程加快,信息资源的数字化、信息传输的网络化改变着人们的生存环境和生活状态,改变了人们获取信息的方式。人们获取信息的主要来源已不再局限于纸质文献,数字资源已成为非常重要的信息来源。

数字化信息资源不仅数量大,其类型亦多种多样,从数字化全文报纸和期刊、电子图书、书目数据库、联机数据库到软件、服务信息、公益信息、动态性信息等;除文献信息外,还有事实、数据、图像、声音、动画等信息;内容涉及科技、政治、经济、商业、新闻、教育、统计、文学、娱乐等各个领域。数字资源数量和种类的无限丰富,以及信息质量的无保证性,直接对我们的学习与信息的获取能力提出了挑战。因此,在网络环境下如何快捷、准确、全面地获取所需的数字资源,并学会分析、利用信息资源已经成为高等院校学生立足于信息社会的一个重要技能。本书是为了适应高等学校教学改革和人才培养目标的需要,帮助高校学生学习和掌握网络环境下数字资源的检索与利用方法而编写的一本通用性和针对性相结合、编排结构全新的信息技术教材。

作为浙江省高等教育重点建设教材,本书较好地体现了理论知识和检索技能并重的原则,保证了整个教学内容的新颖实用、全面系统,是一本数字资源检索与利用的实用教材。

全书共十一章,分为四个部分:第一部分(第1章)主要介绍了数字资源的特点与类型,数字资源的检索原理、检索技术和检索策略等,可为读者学习数字资源及其检索的理论以及掌握数字资源的检索技能奠定扎实的基础。第二部分(第2章~第9章)介绍了中文文献数据库、外文文摘数据库、外文全文数据库、事实数值型数据库以及电子期刊、电子图书、电子报纸等数字资源的检索方法及技巧。该部分从实用的角度考虑,详略结合,力求显示最新的检索方法,以便提高学生的检索技能。第三部分(第10章)介绍了网络学术资源的收集方法,包括常用中外文搜索引擎、网络学术资源导航等。该部分注重实用,力求使读者在最短的时间内掌握 Internet 的应用技能。第四部分(第11章)介绍了信息资源的收集技巧,信息管理工具——E-Reference Tool,论文写作与投稿指南等,指导学生如何搜集、整理、管理和分析利用信息,如何利用信息资源进行论文开题及学位论文写作,如何撰写文献综述和述评,并对国内外论文投稿指南作了介绍。这些内容对提高学生信息素养、培养其运用信息的能力有重要的作用。

本书由章云兰、万跃华、舒炎祥编著,编写分工如下:章云兰编写第1、2、8章,第6.1节,第6.2.1—6.2.3小节。万跃华编写第3—5、7章,第6.2.4小节,第10.1、10.2、11.2节。舒炎祥编写第9章,第6.3、10.3、10.4、11.1、11.3节。

本书得到了浙江省教育厅高教处和浙江省高等学校图书情报工作委员会领导的支

持和帮助；浙江工业大学图书馆馆长何立民教授对全部书稿进行了认真详尽的审阅，提出了许多修改意见，并对全书进行了统稿。在此一并表示衷心的感谢。

此外，在本书的编写过程中，我们参阅了许多教材、论文和网页，并引用了部分论点，限于篇幅，仅列出主要参考文献。借此机会向所有参考文献的作者致以诚挚的谢意。

限于编者的学识水平，加之时间仓促，书中难免有疏漏和不足之处，敬请学界同行、师生和读者批评指正。

目 录

第1章 数字资源的检索原理与检索策略.....	1
引言.....	1
1.1 数字资源及其检索概述.....	2
1.1.1 数字资源的概念与特点.....	2
1.1.2 主要数字资源类型简介.....	3
1.1.3 数字资源检索的发展阶段.....	5
1.1.4 数字资源检索的进展.....	6
1.2 数字资源检索原理.....	8
1.2.1 信息存储.....	9
1.2.2 信息检索.....	10
1.3 数字资源的检索技术.....	11
1.3.1 布尔逻辑检索技术.....	11
1.3.2 截词检索技术.....	12
1.3.3 邻近检索技术.....	13
1.3.4 字段检索技术.....	14
1.4 数字资源检索语言.....	15
1.4.1 分类检索语言.....	15
1.4.2 主题检索语言.....	17
1.4.3 主题语言与体系分类语言的区别.....	19
1.5 数字资源的检索策略.....	20
1.5.1 信息需求分析.....	20
1.5.2 数据库的选择.....	21
1.5.3 选择检索方式.....	21
1.5.4 检索项的确定.....	22
1.5.5 检索式的构造.....	23
第2章 中文文献数据库.....	25
2.1 中国期刊全文数据库.....	25
2.1.1 概述.....	25
2.1.2 数据库的检索途径与方法.....	26
2.1.3 检索结果的输出和下载.....	31
2.2 万方数据资源系统.....	34
2.2.1 概述.....	34
2.2.2 数字化期刊子系统的检索途径与方法.....	35

2.3	中文科技期刊数据库.....	37
2.3.1	概述.....	37
2.3.2	数据库的检索途径与方法.....	38
2.3.3	数据库的辅助检索功能.....	39
2.4	CALIS 数据库.....	41
2.4.1	概述.....	41
2.4.2	CALIS 中文数据资源的检索.....	42
2.5	其他中文文献数据库.....	43
2.5.1	人大复印报刊资料数据库.....	43
2.5.2	国家科技图书文献中心数据库检索.....	45
2.5.3	全国报刊索引数据库.....	49
第 3 章	著名综合性参考数据库.....	52
3.1	科学引文索引.....	52
3.1.1	ISI Web of Knowledge.....	53
3.1.2	ISI Web of Science.....	55
3.1.3	数据库检索.....	56
3.1.4	检索技术.....	68
3.1.5	检索结果.....	69
3.1.6	保存检索历史和创建定题服务.....	73
3.2	工程索引.....	79
3.2.1	检索功能.....	80
3.2.2	检索技术.....	89
3.2.3	检索结果的处理.....	89
3.3	科学技术会议录索引和社会科学与人文科学会议录索引.....	91
3.3.1	ISI Proceeding.....	92
3.3.2	数据库检索.....	92
3.3.3	检索技术.....	97
3.3.4	检索结果.....	97
第 4 章	英文参考数据库.....	102
4.1	科学文摘.....	102
4.1.1	INSPEC 的分类结构和主题词表.....	102
4.1.2	数据库检索.....	105
4.1.3	检索技术.....	113
4.1.4	检索结果.....	114
4.1.5	保存检索历史和创建定题服务.....	118
4.2	化学文摘.....	121
4.2.1	SciFinder Scholar 概述.....	122
4.2.2	SciFinder Scholar 数据库检索.....	124
4.2.3	CA on CD 光盘数据库.....	131

4.3	生物学信息数据库	136
4.3.1	数据库检索	136
4.3.2	检索技术	144
4.3.3	检索结果	145
4.4	OCLC FirstSearch 系统数据库	148
4.4.1	FirstSearch 系统	148
4.4.2	FirstSearch 数据库	149
4.4.3	数据库检索	152
4.4.4	检索技术	155
4.4.5	检索结果	156
4.4.6	管理模块	157
第 5 章	全文数据库	158
5.1	全文数据库概述	158
5.2	ProQuest 系统全文数据库	160
5.2.1	ProQuest 系统主要的数据库	160
5.2.2	数据库检索	164
5.3	EBSCOhost 系统全文数据库	164
5.3.1	EBSCOhost 系统主要的数据库	165
5.3.2	数据库检索	167
5.4	LexisNexis 系统全文数据库	178
5.4.1	LexisNexis Academic 学术大全数据库	178
5.4.2	Lexis.com 数据库	183
第 6 章	电子图书、电子报纸及其使用	184
6.1	电子图书和报纸概述	184
6.1.1	电子图书	184
6.1.2	电子报纸	186
6.2	数字图书馆及其电子图书服务	187
6.2.1	书生之家数字图书馆	187
6.2.2	超星数字图书馆	190
6.2.3	其他中文数字图书馆及电子图书服务	192
6.2.4	外文电子图书	193
6.3	电子报纸及其利用	200
6.3.1	电子报纸网站	201
6.3.2	报纸类数据库	202
第 7 章	电子期刊及其使用	206
7.1	电子期刊概述	206
7.2	著名出版社的英文电子期刊	207
7.2.1	Elsevier Science 出版社的电子期刊	207
7.2.2	Kluwer Academic 出版社的电子期刊	212

7.2.3	Springer 出版社的电子期刊	216
7.2.4	John Wiley 出版社电子期刊	219
7.2.5	World Scientific Publishingy 出版社的电子期刊	223
7.2.6	Nature 出版社的电子期刊	224
7.2.7	Blackwell 出版社的电子期刊	225
7.2.8	Oxford 大学出版社的电子期刊	226
7.3	著名学会出版的英文电子期刊	226
7.3.1	美国化学学会 (ACS) 电子期刊	226
7.3.2	英国皇家化学学会 (RSC) 电子期刊	229
7.3.3	美国物理研究所 (AIP) 和美国物理学会 (APS) 的电子期刊	229
7.3.4	英国皇家物理学会 (IOP) 的电子期刊	230
7.3.5	美国计算机学会 (ACM) 电子期刊	231
7.3.6	美国电气电子工程师学会 (IEEE) 和英国电气工程师学会 (IEE) 的电子期刊	231
7.3.7	美国机械工程师学会 (ASME) 电子期刊	232
7.3.8	美国土木工程协会 (ASCE) 电子期刊	232
7.3.9	科学在线 (Science Online) 电子期刊	233
7.3.10	JSTOR 英文过期电子期刊	234
7.4	免费电子期刊	236
7.4.1	斯坦福大学 HighWire Press 出版社的电子期刊	236
7.4.2	开放使用学术期刊	237
第 8 章	特种文献及其检索	240
8.1	专利文献及其检索	240
8.1.1	专利与专利文献概述	240
8.1.2	中国专利信息的检索	241
8.1.3	国外专利信息的检索	246
8.2	会议文献及其检索	254
8.2.1	会议文献概述	254
8.2.2	国内会议文献的检索	255
8.2.3	国外会议文献的检索	256
8.3	学位论文及其检索	257
8.3.1	学位论文概述	257
8.3.2	美国博硕士学位论文数据库 (PQDD)	258
8.3.3	中国学位论文数据库	262
8.3.4	其他学位论文数据库	265
8.4	科技报告及其检索	266
8.4.1	科技报告概述	266
8.4.2	NTIS 数据库检索	267
8.4.3	AD 报告数据库检索	268
8.4.4	DOE 报告数据库检索	269

8.4.5	NTRS 数据库的检索	269
8.4.6	其他科技报告数据库的检索	270
8.5	标准文献及其检索	270
8.5.1	概述	270
8.5.2	中国标准的检索	271
8.5.3	国外标准信息检索	273
第 9 章	网络学术资源的收集	277
9.1	搜索引擎的比较与应用	277
9.1.1	搜索引擎概述	277
9.1.2	常用中文搜索引擎及特点	279
9.1.3	常用英文搜索引擎及特点	287
9.1.4	搜索引擎的发展趋势	291
9.2	网络学术资源导航及其应用	292
9.2.1	概述	292
9.2.2	网络学术资源的内容	293
9.2.3	国内外网络学术资源导航系统介绍	294
9.2.4	如何在学术研究中利用网络资源导航	295
9.3	网络学术资源的其他收集方法	296
第 10 章	事实和数值型数据库	298
10.1	事实和数值型数据库概述	298
10.2	英文事实和数值型数据库	299
10.2.1	DEALOG 商情数据库	299
10.2.2	人类生物基因组数据库	299
10.2.3	贝尔斯坦/盖墨林化学数据库	301
10.2.4	ISI 化学数据库	305
10.2.5	Gale 出版集团的参考性资料数据库	311
10.2.6	LexisNexis 参考资料数据库	313
10.2.7	《不列颠百科全书》网络版	314
10.3	中文事实数值数据库	315
10.3.1	万方数据资源系统之事实和数值数据库	315
10.3.2	中国科学院“科学数据库”	317
10.3.3	中国资讯行高校财经数据库	318
10.3.4	新华社多媒体数据库高等教育版	320
10.3.5	中经网: 教育版	323
10.3.6	国研网	325
10.4	其他事实与数值型信息检索	327
10.4.1	网上事实与数值型信息资源的评估	327
10.4.2	网上事实与数值型资源站点选介	328

第 11 章 信息资源的分析利用、组织与论文写作	335
11.1 信息资源的收集技巧.....	335
11.1.1 信息资源的其他收集方法	335
11.1.2 信息资源的收集技巧	337
11.2 个人信息管理工具——E-Reference Tool	339
11.2.1 个人信息管理工具概述	339
11.2.2 EndNote 收集、整理、管理和引用参考文献	341
11.2.3 Reference Manager 收集、整理、管理和引用参考文献	350
11.3 学位论文写作与投稿指南	354
11.3.1 学位论文的特点、写作步骤和格式	354
11.3.2 学位论文的开题查询	360
11.3.3 文献综述的特点和写作方法	362
11.3.4 文献述评的特点和写作方法	363
11.3.5 文摘的类型与编写方法	364
11.3.6 国内外论文投稿指南	366

第1章 数字资源的检索原理与检索策略

引言

当今社会已进入信息化时代。信息时代,就是以信息资源为主要资源,以信息资源的获取和产生为核心,来促进社会和经济发展的时代。在这一时代里,信息已经成为个人、社会群体和整个社会发展的极其重要的资源,人们需要依赖信息而生存发展。特别是在科学研究中,无论是一个课题的选题、设计、实验,还是成果鉴定,每一步都离不开信息。有资料表明:科研工作者在科研的过程中要用 80% 的时间来获取信息,而只用 20% 的时间来加工和产生新的成果。因为只有通过信息的查阅,才能明确研究的可行性,并把自己的研究工作建立在一个较高的起点上。相反,如果在科研中,不重视信息的查阅,没有做好继承和借鉴工作,则容易重复研究,浪费大量的人力、物力和财力。据统计,美国每年由于重复研究所造成的损失,约占全年研究经费的 38%,达 20 亿美元之巨。日本有关化学化工方面的研究课题与国外重复的,大学占 40%、民间占 47%、国家研究机构占 40%,平均重复率在 40% 以上;我国的重复率则更高。

随着信息社会文献载体的发展、进化以及信息技术和因特网的飞速发展,大量的数字化信息资源通过信息网络快速而平等地摆到了每一个人的面前。人们获取信息的主要渠道已不再是纸质文献,数字资源已成为重要的信息来源。在整个社会的信息消费体系中,数字资源的消费指数正在成倍地增长,并逐步动摇着印刷型信息长期以来在信息消费体系中的主导地位。

海量的数字资源和“信息爆炸”的网络,极大地拓宽了人们获取信息的空间,但同时也给人们检索与利用信息带来了新的问题和矛盾。

首先是数字资源总量的剧增与个人信息获取能力有限之间的矛盾。以 Internet 上的数字资源为例,2003 年 7 月,CNNIC(中国互联网信息中心)的报告显示,全球因特网使用者现有 5 亿人,平均每分钟就有 97 万封电子邮件被发送,以 WWW 方式提供的数字资源平均每过 53 天就翻一番,全球现已有超过 30 亿网页及 2000 万个网站。如此大的信息量对个人来讲是个天文数字,因为每个人的时间、精力与信息的无限是不可比拟的,于是数字资源总量的剧增与个人信息获取能力有限的矛盾便越来越突出。

其次是数字资源生产、发布的任意性、自由性与信息质量无保证之间的矛盾。数字资源的生产和发布不再仅限于官方或正式的出版机构,任何人都可以成为网上信息的发布者,因而对信息缺乏控制和管理,没有认证和审核,使得数字资源质量良莠不齐,各种干扰信息、噪音信息充斥信息共享空间,影响了人们对数字资源的获取和信息质量的评估。

再次是数字资源种类的多样性与人们信息检索能力之间的矛盾。现代信息处理、传播与交流技术为人们提供了多种多样的数字资源类型,信息的形式不再是视觉和静态的形式,而是文本、数据、图形、图像、音频和视频等多媒体和动态形式,这些信息均由

功能强大的计算机软件系统来组织和管理,检索这类信息资源也需要一定的技能。不同国家、不同地区、不同行业、不同人群、不同年龄的人们之间对信息的了解、对信息检索技术的应用程度都存在差别,这就造成了所谓的“数字鸿沟”的存在,不能跨越这道鸿沟的人最终将会被弱化和边缘化。

以上矛盾的客观存在阻碍了数字资源的传播、交流和共享利用,影响了获取信息的效率。因此,网络环境下数字资源的检索与利用能力就成为人们迫切需要解决的问题。

编写本书的目的在于通过对数字资源及其检索方法与途径的介绍,解决人们在网络环境下查找信息所面临的一系列问题,提高大学生对数字资源的检索、选择、吸收、整理和重组信息的能力。

1.1 数字资源及其检索概述

1.1.1 数字资源的概念与特点

1. 数字资源的概念

数字资源是数字化的信息资源或数字化文献,是指可通过计算机本地或远程读取、使用,以数字形式存放在光、磁载体上,以电信号、光信号的形式传输的图像、文字、声音、视频等信息资源。

数字资源的范围非常广泛,其类型繁杂,形式多样,根据不同的划分标准可分为不同的类型。

1) 按生产途径和发布范围的不同,可分为商用资源、网络公开学术资源、特色资源和其他资源。

商用资源是由正式出版机构或出版商/数据库商出版发行的数字资源,在数字学术信息资源中所占比例最大,包括各类数据库和电子期刊、电子图书等。

网络公开学术资源是完全面向公众开放使用的数字资源,包括各种学术团体、行业协会、政府机构、商业部门、教育机构等在網上正式发布的网页及其信息。

特色资源,主要指基于各教育机构、政府机关、图书馆的一些特色收藏制作,在一定范围内分不同层次发行的数字资源。

其他资源,如FTP资源、新闻组、BBS、电子邮件等。

2) 按储存的物理地点的不同,数字信息资源又可分为现实资源和虚拟资源。

现实资源是指置放于本地本部门的数字化文献,有光盘、磁盘、磁带等多种载体形态。包括出版商出版的电子出版物、本单位制作的数字化文献和从网络上下载到本地的信息资源。

虚拟资源又称网络资源或联机资源,是指通过本单位的计算机系统及通信网络才能获取的数字化信息,是置放于异地的数字化信息。

3) 按信息源的不同,可分为数据库、电子期刊、电子图书、电子报纸、联机馆藏目录和网络资源等。

2. 数字资源的特点

1) 数字资源是以现代信息技术为记录手段而形成的一种数字化的信息,这种数字化的信息以机读数据的形式存在,既可以在计算机内被高速处理,又可以通过通信网络进行远距离传送,不受时间和空间的限制,这就使全球信息资源共享成为可能。随着网络的进一步扩大与应用范围的拓宽,数字化的信息资源将成为信息资源的最终转化方式。

2) 数字资源是以磁性材料或光学材料为存储介质的,信息资源由纸张上的文字变成磁性介质上的电磁信号或光介质上的光信息,从而使信息的存储传递和查询更加方便,而且所存储的信息密度高,容量大,占用空间少,检索速度快,检索功能齐全,能提供题名、著者、主题词、关键词、号码、年代、出处等多种检索途径。

3) 数字资源包含的信息类型多种多样,从电子报刊、电子工具书、商业信息、新闻报道、书目数据库、文献信息索引到统计数据、图表、电子地图等;既可以是文字、图表等静态信息,也可以是集图、文、声、像于一体的动态多媒体信息;并且各种类型的数据又可借助计算机实现任意的组合编辑,具有打印输出及磁盘输出等多种输出方式。这就使得信息的组织方式发生了巨大的变化,不仅以知识和信息为基本单元,而且充分展示这些单元间的逻辑关系。

4) 数据结构具有通用性、开放性和标准化的特点。在网络环境下,数据可以被多人同时访问,是一种共享性的信息资源,这就使得数字资源更易于实现资源的扩充。

5) 网络环境下,数字资源传递和反馈快速灵敏,具有动态性的实时性的特点。数字资源在网络的流动非常迅速,电子流取代了纸张和邮政的物流,加上无线电和卫星通信技术的充分运用,任何数字信息源,只要上传到网上,就能够在短短的数秒内传递到世界的每一个角落。

6) 数字资源的组织形式从顺序的、线性的方式转变为直接的、网状的组织形式;由传统的顺序排列,利用数字化存储技术,发展到超文本技术。数字资源的组织形式的转变使数字资源可按自身的逻辑关系组成相互联系的网络结构,用户不必关注所查信息存在何处,只须由检索工具完成。

1.1.2 主要数字资源类型简介

数字资源的范围非常广泛,类型有多种多样,在此就本书将要重点讲述的数字资源作简要介绍。

1. 参考数据库

参考数据库有时又称为书目数据库,是指包含各种数据、信息或知识的原始来源和属性的数据库。数据库中的记录是通过对数据、信息或知识的再加工和过滤,如编目、索引、摘要、分类等,然后形成的。到目前为止,参考数据库主要是针对印刷型出版物开发的,它的作用是为用户指出了获取原始信息的线索,目的是指引用户能够快速、全面地鉴别和找到相关信息。

参考数据库主要包括目录数据库、文摘数据库和索引数据库。目录数据库主要是

针对图书进行内容和存储地址的报道与揭示的, 又称机读目录数据库, 其数据内容详细, 除描述标题、作者、出版项等书目信息外, 还提供用户索取原始信息的馆藏信息。文摘和索引数据库则相对期刊论文、会议论文、专利文献、学位论文等进行内容和属性的认识与加工。它提供确定的文献来源信息, 供人们检索, 但一般不提供原始文献的馆藏信息。

2. 全文数据库

全文数据库是指收录有原始文献全文的数据库, 它将一个完整的信息源的全部内容转化为计算机可以识别、处理的信息单元而形成的数据集合。不仅存储了信息, 而且还有对全文数据进行词、字、段落等更深层次的编辑、加工的功能。主要以期刊论文、会议文献、学位论文、研究报告、法律条文、商业信息等等为主。

3. 事实、数值型数据库

事实、数值型数据库指包含大量数据、事实、概念, 直接提供原始资料的数据库。

1) 事实型数据库以各种有检索和利用价值的事实信息为存储对象, 库内记录各种名词术语、有关学科、机构、名人等方面的简要情况, 其数据资料来源于百科全书、字典、人名录、公司名录等。

2) 数值型数据库以各种调查数据或统计数据为存储对象, 库内记录各种有价值的数值、有关的运算公式和规则等信息, 如各种物理常数、物质分子量、元件参量、市场行情调查数据、股票证券交易数据、有关部门的统计数据等。这类数据库均以数值形式回答用户提问, 在回答之前, 系统通常还需按用户要求作某些必要的计算。

4. 电子图书、电子期刊和电子报纸

1) 电子图书 (Electronic Book, 简称 eBook): 又称数字图书, 是纸质图书 pBook (Paper Book) 的对应物。它是一种以数字形式在因特网上出版、发行, 读者可通过阅读终端进行下载的数字化书籍。

2) 电子期刊: 正如同印刷型期刊在传统的学术资源中占据很重要的地位一样, 电子期刊在数字信息资源中也独树一帜, 正在发挥越来越大的作用。目前在网上的纯电子版期刊和印刷版期刊的电子版已有两万种左右, 这些期刊具有出版周期短, 可以检索和重复下载全文, 图像与文本结合, 包含有多媒体及其他类型动态信息, 具备超链接功能等特点, 越来越受到用户的欢迎。

3) 电子报纸: 就是在排、印、投递等方面基本上实现了电子化的报纸, 主要有印刷型报纸的电子版和纯电子报纸两种类型, 目前主要以光盘版和网络版两种形式发行。

5. 网络资源指南/搜索引擎

网络资源极为丰富, 在地理上涵盖天南地北, 在内容上更是包罗万象, 因此必须依赖一些工具, 才能有效地检索、浏览、传输和利用, 其中最好的方法就是使用各种网络资源指南以及搜索引擎进行检索。

1.1.3 数字资源检索的发展阶段

数字资源检索起源于 20 世纪 50 年代初。1954 年美国海军兵器中心图书馆利用 IBM701 机开发计算机信息检索系统,它标志着数字资源检索阶段的开始。目前,数字资源检索已成为用户获取信息的主要方式,出现了联机信息检索、光盘信息检索和网络信息检索等多元并存的格局。综观数字资源检索的发展,可以将其划分为脱机检索、联机检索、光盘检索和网络检索等 4 个发展阶段。

1. 脱机检索阶段

20 世纪 50—60 年代是脱机检索的试验和实用化阶段,批处理检索方式是这个阶段信息检索的主要方式。著名的脱机检索系统有美国国家医学图书馆的 MEDLARS,美国化学文摘社发行的《化学题录》机读磁带版等。

在脱机检索中,必须事先把众多用户的各种信息要求编成“用户提问档”格式,以机读形式存储在磁带上,定期地(一个月或半个月)检索数据库新增加的内容,然后把命中的文献信息发给用户。它由专职检索员统一处理,而且不提供任何实际浏览的可能性,用户不接触机器,所以它是一项“机会检索”。这是一种早期的数字资源检索形式,其数据存取与数据通信能力都比较差。

2. 联机检索阶段

20 世纪 60—80 年代是联机检索试验和实用化阶段。1960 年美国麻省理工学院开始实施有关联机检索系统设计的“技术情报计划”,系统发展公司也在它开发的全文检索系统上进行了首次联机检索演示,该公司后来研制成功的联机信息检索软件 OBIT 是联机检索阶段的正式开始。

联机检索是指用户利用终端设备,通过通信网络与计算机检索系统连接,远距离进行“人机对话”,从检索系统的数据库中检索出所需信息的过程。各联机检索系统的检索指令和检索运算符各不相同,一般需由具备专门检索技术的专业检索员来检索。目前世界上已有 600 多个商用联机检索服务系统,其中以 DIALOG、STN、OCLC、BRS、ORBIT、ESA-IRS 等系统最为著名。

传统的联机检索系统绝大多数是集中式管理的,它的主要优点是系统的安全性以及具有在存储设备上处理大量数据的功能。但也存在许多缺点,如主机的负担很重,对主机的依赖性过高,网络的扩展性较为困难等。

随着 Internet 的出现和 ISP 的高速增长,传统的联机信息检索业开始遇到极大的挑战。因此,近年来世界上许多联机检索系统如 DIALOG、STN、OCLC 等都在因特网上建立了 Client-Server 和 Telnet 的应用环境,成为 Internet 的一部分。联机检索系统上网后,在服务项目与内容方面除了保持原有的特色以外,还吸收了 WWW 服务器信息服务方面的长处,增加了许多服务项目和内容,如消息、网络新闻组等。此外,在 Internet 上,联机检索系统用户界面获得了极大的改善,使用同一浏览器以及图形用户界面,系统的可操作性大大加强,使更多的人接受了这类系统。

3. 光盘检索阶段

光盘检索阶段始于 20 世纪 80 年代中期, 1985 年世界上第一个 CD-ROM 数据库 BIBIIOFICE (美国国会图书馆机读目录) 问世, 是光盘检索系统实用化的标志。

光盘检索是以光盘为存储介质, 利用光盘驱动器和计算机实现对光盘数据库的读取和检索的系统。这种检索系统的结构比较简单, 任何一台安装了光驱的计算机, 只要放上光盘数据库即可成为光盘数据库检索系统。由于光盘数据库的制作技术较为成熟和简便, 所以光盘数据库的数量激增, 内容广泛, 成为深受用户喜爱的一种数字信息源。在单机光盘数据库检索系统的基础上, 20 世纪 90 年代又开发了光盘塔和光盘网络软件, 使光盘数据库检索系统实现了局域网范围内共享。此外, 在 Internet 日益成为信息服务主流渠道的今天, 许多传统的光盘数据库生产商也开始建立起自己的信息网站, 在继续发行光盘版数据库的同时, 将光盘数据库的内容放置在网上, 实现了光盘检索的网络化, 出现了跨网段、跨广域网的多模式、多功能的检索。

4. 网络检索阶段

网络信息检索开始于 20 世纪 90 年代初。1991 年思维机等公司、明尼苏达大学、欧洲高能粒子协会分别推出了因特网上的检索工具 WAIS、Gopher 和 WWW。

目前, WWW 因其集文本、图像、声音等多媒体信息于一体的巨大优点, 已占信息服务的主导地位, 基于 Web 的搜索引擎已成为最重要的信息检索工具。

在因特网中, 主要使用的是分布式的计算机信息检索系统。它与高性能的并行计算机及大型的数据库管理系统共同管理、调度、分配网中的通道和信息资源。网络中的数据库并不是像联机检索系统那样集中于检索中心而是分布在网络中的各个主机。这种分布式数据可以实现分布式存取、分布式存储和分布式处理, 因此极大地提高了网络的安全性, 并具备模块性、并行性、共享性的主要特点。

以往网络上的信息质量不够可靠, 但是近年来, 许多传统的信息服务领域纷纷向因特网进军, 将自己权威的信息通过独立的网站向全球及时发布, 从而大大提高了因特网的信息服务质量。网上的信息一般都通过搜索引擎提供多途径检索, 检索界面十分友好, 无须掌握复杂的检索技巧, 使数字资源检索突破了传统科研学术领域的限制, 更深入地融合到了人们的日常生活之中。

1.1.4 数字资源检索的进展

自 20 世纪 90 年代以来, 以 Internet 为核心联合起来的全球计算机网络使传统相对集中和规范的文献数据库 (也称参考数据库) 及其检索系统面临挑战。因为 Internet 的发展, 传统检索的中介代理的功能正在逐步减弱, 成千上万的人都将成为网络系统的最终用户。网络系统中存储的内容除了原来的二次信息外, 已出现越来越多的全文数据、事实数据、数值、图像和其他多媒体信息资源。信息的可获性、传递速度大大提高和增强, 跨文件、跨文档和跨数据库以及在多媒体数据库中自由浏览已成为现实。这些变化都是以现代信息技术及网络技术发展为前提的。