

数字信息资源的 检索与利用

主编 肖 珑
编著 肖 珑
张春红 苏玉华
李浩凌 张宇红 廖三三

北京大学出版社



数字信息资源的检索与利用

主 编 肖 琰
编 著 肖 琰 张春红 苏玉华
李浩凌 张宇红 廖三三

北京大学出版社
北 京

内 容 提 要

互联网时代,几乎每个人都知道可以到网上寻找自己想要的信息。但网络资源种类繁多,千变万化,对于查找学术信息的科研、教学人员、学者专家、学生来说,要想在短时间内满足需求,做到事半功倍,就必须首先掌握相关的检索和导航工具。

本书介绍了300余种国内外著名的数字学术资源/相关检索工具的内容及其查询方法,涉及300余个文摘索引/原始一次文献数据库、近2万种电子期刊、500多种电子报纸、十几万种电子图书,以及会议论文、学位论文等,中英文资源各占一半,同时包括搜索引擎、网络学术资源导航等其他导航工具。全书体系完整,图文并茂,为读者准确、快速、全面地查找网络学术信息提供了颇为有效的帮助。

本书是用户查找网络学术资源的工具书,也可作为大专院校、科研机构开展信息素质教育/检索培训的教材,并供相关人员参考。

图书在版编目(CIP)数据

数字信息资源的检索与利用/肖珑等编著. —北京:北京大学出版社,2003.7

ISBN 7-301-06369-5

I. 数… II. 肖… III. 计算机应用—情报检索 IV. G252.7

中国版本图书馆CIP数据核字(2003)第052345号

书 名: 数字信息资源的检索与利用

著作责任者: 肖 珑 主编

责任编辑: 沈承凤

标准书号: ISBN 7-301-06369-5/Z·0080

出 版 者: 北京大学出版社

地 址: 北京市海淀区中关村 北京大学校内 100871

网 址: <http://cbs.pku.edu.cn> 电子信箱: zpup@pup.pku.edu.cn

电 话: 邮购部 62752015 发行部 62750672 编辑部 62752038

排 版 者: 兴盛达打字服务社 82715400

发 行 者: 北京大学出版社

印 刷 者: 北京大学印刷厂

787毫米×1092毫米 16开本 26印张 649千字

2003年7月第1版 2003年7月第1次印刷

定 价: 38.00元

题 辞

随着信息化建设的推进，数字信息资源迅速增长，如何方便有效地利用这些宝贵信息成了重要问题。肖珑同志主编的这本书将有助于读者了解数字信息资源、相关的检索工具，以及如何有效地查阅这些信息。希望此书能对读者有所帮助。

王 选

2003. 4. 22

前 言

在北大图书馆咨询台值班的时候,总会碰到读者询问类似的问题:“老师,怎么在网上找到这篇文献?”,或者“老师,某某数据库怎么检索?”,等等,算起来,这些问题每个月至少可以达到数以百计。综合这些提问,不外乎两类问题:

- (1) 怎样在网上找到相关信息,有哪些工具可以利用?
- (2) 这些网络工具或者数字资源如何使用?

针对这些需求,我们先是在北大图书馆开设了名为“一小时讲座”的系列讲座,为读者集中系统地讲解这些问题,稍后我们又开设了名为“电子资源的检索和利用”的全校性公共选修课。三年来,累计有近两万人次聆听或选修了我们的讲座和授课。为了讲座和授课的需要,我们同时撰写了相应内容的大量讲稿。随着讲义篇幅内容的不断增多,我们不由得产生了一个想法:写一本指南性的工具书,告诉读者,网上有哪些数字学术资源可以利用,这些资源使用的方法是什么,在科研、教学和学习中如何综合利用这些资源。

许多同行也对我们说,北大图书馆购买的网上学术资源在高校乃至全国的图书馆中是最多的,你们是不是可以把这些资源写一写,介绍给我们,帮助我们进行本校的数字资源建设和为读者讲授数字资源的检索。于是,我们在授课和讲座的讲义的基础上,开始动手写这本书。

本书的写作历时两年。其间,随着资源的日益增加和数据库的情况变化,书的内容不断地丰富和修改,篇幅也逐渐加大,从计划的30万字变成了最后的60多万字。现在,书稿终于完成,衷心希望本书能够达到我们预想的目的,为用户和同行利用网上数字学术资源提供参考和指南。

全书内容共十章,分为四个部分:

第一章为方法论,对数字信息资源及其结构、基本检索方法进行了概述。

第二章至第九章,按照第一章的方法论内容,介绍了约70家出版商/数据库商/学术机构的300余个数字学术资源的内容及其使用方法,涉及200多个数据库、近2万种全文电子期刊、500多种电子报纸、十几万种电子图书,其中中英文资源大约各占一半;此外,对搜索引擎、网络学术资源导航、FTP等其他导航工具也做了介绍。

第十章为上述资源的综合利用,列举了若干实例。

最后部分为资源索引,方便读者从字顺、学科、出版商/数据库商/学术机构等途径查找资源。

每章后面均附有该章的参考文献。

在第二章至第九章中,对不同的数字资源按照参考数据库、全文数据库、事实数据库、电子期刊、电子图书、电子报纸和其他数字学术资源等类型分别进行了介绍,同时尽量采取了相同体例,每种资源的介绍大致包括:

(1) 数据库内容:数据库的中英文名称,出版商/数据库商/学术机构名称,发展历史,收录出版物,涵盖学科,收录年限,更新频率,网址,特点,与相关数据库或不同版本之间的比较等。

(2) 检索方法:

- ① 检索功能:简单检索,复杂检索,自然语言检索,图像检索,浏览,索引,可检索字段等;
- ② 检索技术:布尔逻辑,组配检索,截词算法,位置算符,词根运算,嵌套(优先)运算,大小写敏感,相关检索,引文检索,禁用词表等;
- ③ 检索结果:检索结果格式、标记方法及下载方法等;
- ④ 相关用户服务:如相关帮助文件/培训教程,检索史记录,个性化服务,最新文献报道服务,文献传递服务,相关参考工具等。

(3) 用户检索界面及检索结果的图片及相关说明。

此外,对有些历史悠久的权威数据库,为方便用户使用,还简要介绍了其印刷版本的结构、体例和使用方法;对有些跨类型、跨章节介绍的数据库,在文中以参见形式进行了介绍。

在此需要说明的是,书中所介绍的资源,有些是需要购买访问权限的;有些网址可能会发生迁移变化;所采用的相关数据大多截止到2002年12月底,之后发生的变化并没有记录在内;这些方面读者如有疑问,可随时与我们联系。

全书的写作分工如下:

肖珑:第一章、第五章、第七章、第九章第一节,此外还负责全书的策划组织、体例制订和最后的统稿。

张春红:第二章、第八章、第九章第二节、第十章第二节以及索引的汇总、校对;

苏玉华:第四章、第十章第三节;

李浩凌:第六章、第十章第一节;

张宇红:第三章;

廖三三:第九章第一节(与肖珑合写)、第九章第三节;

北京大学数字图书馆研究所的研究生胡良霖同学参加了第九章第一节、第三节的修订工作。

在过去的两年中,由于数字信息资源和数字化服务在北京大学图书馆的迅速发展,使得所有参加本书编撰的同事,都处在一种高度繁忙的工作状态中,本书几乎全部是大家利用业余时间撰写的。尽管如此,为了充分满足读者的需求,给用户是一本高质量的工具书,每个人都以极其认真、细致的态度撰写自己负责的部分,并不厌其烦地反复修改。在此,我要衷心地感谢各位参加编写工作的同事,对他们的执着精神和认真负责的工作态度表示敬意。

感谢中国科学院院士、北京大学王选教授,他在百忙中看了书稿,并欣然为本书题辞,肯定了本书的价值和作用。王选教授的题辞对我们是莫大的鼓舞和鞭策。

本书在编写过程中,得到了北京大学图书馆戴龙基馆长、中国高等教育文献保障系统(CALIS)管理中心陈凌副主任、北京大学教材办公室张积老师、北京大学出版社沈承凤、郭力老师以及北京大学图书馆信息咨询部的其他各位同仁的大力支持和帮助,在此一并表示感谢。

本书错谬之处,欢迎读者批评指正。

肖 珑

2003年4月

目 录

第一章 数字信息资源及其检索概述	(1)
引言	(1)
第一节 数字信息资源的概念与类型	(3)
1.1 数字信息资源的产生与发展	(4)
1.2 数字信息资源的类型	(8)
第二节 数字信息资源的检索	(10)
2.1 数字信息资源的检索系统	(11)
2.2 与数字资源检索相关的网络基础知识	(14)
2.3 联机数据库检索	(18)
2.4 光盘数据库检索	(19)
2.5 网络数据库检索	(21)
2.6 检索语言	(22)
第三节 数字信息资源的检索方法与检索技术	(28)
3.1 检索方法	(29)
3.2 检索功能	(31)
3.3 检索技术	(35)
参考文献	(39)
第二章 参考数据库——著名综合检索工具介绍	(40)
第一节 参考数据库概述	(40)
1.1 参考数据库	(40)
1.2 著名综合性参考数据库——四大权威检索数据库的形成与发展	(41)
第二节 引文索引与“科学引文索引”(SCI)	(42)
2.1 引文索引	(42)
2.2 SCI的发展历史、出版类型、内容和特点	(43)
2.3 SCI印刷版的结构与编排	(45)
2.4 SCI光盘数据库检索(Windows版)	(46)
2.5 SCI网络数据库检索——WOS和WOK系统中的SCIE检索	(50)
2.6 SCI的联机检索服务方式介绍	(56)
第三节 社会科学引文索引(SSCI)和艺术与人文科学引文索引(A&HCI)	(57)
3.1 数据库介绍	(57)
3.2 SSCI和A&HCI印刷版的结构与编排	(58)
3.3 SSCI和A&HCI光盘数据库检索(Windows版)	(58)
附: 期刊引用报告(JCR)和JCRWeb	(59)
第四节 工程索引(Ei)	(61)
4.1 数据库介绍	(61)
4.2 印刷版的结构与编排	(63)
4.3 Ei光盘数据库检索	(64)
4.4 Ei网络数据库检索	(69)
第五节 科学技术会议录索引(ISTP)和社会科学与人文科学会议录索引(ISSHP)	(73)
5.1 数据库介绍	(73)

5.2 使用方法	(73)
参考文献	(78)
第三章 常用中文参考数据库	(79)
第一节 中国科学引文数据库(1989~)	(79)
1.1 数据库概况	(79)
1.2 数据库的特点和作用	(80)
1.3 数据库检索	(80)
1.4 检索结果	(82)
第二节 中文科技期刊数据库(1989~)	(84)
2.1 数据库概况	(84)
2.2 光盘版数据库的特点及其检索	(85)
2.3 网络版数据库的特点及其检索	(88)
第三节 万方数据资源系统	(91)
3.1 万方数据资源系统概述	(91)
3.2 万方数据资源系统参考数据库	(92)
3.3 数据库检索	(94)
3.4 检索技术	(97)
3.5 检索结果	(97)
3.6 光盘数据库与网络版数据库的比较	(98)
第四节 中国专利数据库(1985~)	(99)
4.1 数据库概况	(99)
4.2 数据库检索	(100)
4.3 检索结果	(102)
第五节 中文社会科学引文索引(1998~)	(102)
5.1 数据库概况	(102)
5.2 数据库特点	(103)
5.3 数据库检索	(103)
5.4 检索技术	(105)
5.5 检索结果	(105)
第六节 全国报刊索引数据库(社科版、科技版)	(106)
6.1 《全国报刊索引》印刷版的编排及使用	(106)
6.2 数据库概况	(106)
6.3 数据库应用实例	(108)
第七节 中国人民大学书报资料中心复印报刊资料索引总汇	(109)
7.1 《报刊资料索引》印刷版概况	(109)
7.2 数据库特点	(110)
7.3 光盘版数据库	(110)
7.4 网络版数据库	(112)
第八节 CALIS 数据库	(115)
8.1 数据库介绍	(115)
8.2 数据库检索	(116)
8.3 检索技术	(117)
8.4 检索结果	(117)
第九节 国家科技图书文献中心数据库	(118)
9.1 数据库内容及使用	(119)
9.2 NSTL 服务介绍	(121)

参考文献	(122)
第四章 常用英文参考数据库	(123)
第一节 科学文摘(INSPEC)	(123)
1.1 概况	(123)
1.2 INSPEC 的分类结构和主题词表	(125)
1.3 《科学文摘》印刷版简介	(127)
1.4 数据库检索	(128)
1.5 检索技术	(132)
1.6 检索结果	(133)
1.7 INSPEC 的其他电子产品	(135)
第二节 化学文摘(CA)	(135)
2.1 概况	(136)
2.2 CA 印刷版简介	(139)
2.3 CA 光盘版检索	(140)
2.4 网络版数据库——SciFinder	(147)
第三节 生物学文摘(BA)和生物学信息数据库(BP)	(148)
3.1 BA 印刷版简介	(149)
3.2 BA 光盘数据库(BA-WINSPIRS)	(149)
附:使用 WINSPIRS 检索系统的其他数据库	(156)
3.3 生物学信息数据库(BP)	(157)
第四节 OCLC FirstSearch 系统数据库	(163)
4.1 FirstSearch 系统	(163)
4.2 FirstSearch 数据库	(164)
4.3 检索功能	(166)
4.4 检索技术	(169)
4.5 检索结果	(170)
4.6 管理模块	(171)
第五节 剑桥科学文摘数据库(CSA)	(171)
5.1 数据库内容	(172)
5.2 检索功能	(175)
5.3 检索技术	(178)
5.4 检索结果	(179)
5.5 帮助功能(Help)	(182)
5.6 CSA(IDS 系统)的管理模块	(182)
第六节 德温特创新索引(DII)及其他专利数据库	(182)
6.1 概述	(182)
6.2 德温特创新索引	(183)
6.3 其他重要的专利数据库	(190)
第七节 最新目次数据库(CCC)	(190)
7.1 数据库内容	(190)
7.2 数据库检索	(191)
参考文献	(196)
第五章 全文数据库与全文服务	(197)
第一节 全文数据库概述	(197)
第二节 ProQuest 系统全文数据库	(199)
2.1 数据库内容	(199)

2.2	数据库检索	(202)
2.3	系统管理服务	(210)
第三节	EBSCOhost 系统全文数据库	(210)
3.1	数据库内容	(210)
3.2	数据库检索	(212)
3.3	系统管理服务	(216)
第四节	LEXIS-NEXIS 系统全文数据库	(216)
4.1	学术大全数据库(LEXIS-NEXIS Academic, 简称 Academic)	(216)
4.2	Lexis.com 数据库	(221)
第五节	其他英文全文数据库	(225)
5.1	IEEE/IEE 电子图书馆数据库(IEEE/IEE Electronic Library, 简称 IEL)	(225)
5.2	Gale 全文数据库	(228)
5.3	OCLC 全文数据库	(228)
第六节	中文全文数据库	(229)
6.1	中国人民大学书报资料中心复印报刊资料全文数据库	(229)
6.2	“CGRS 全文检索系统”内的其他中文全文数据库	(230)
6.3	其他中文全文数据库	(230)
第七节	互联网上的全文服务	(231)
7.1	全文链接服务(link to full-text)	(231)
7.2	原文传递服务(document delivery)	(232)
	参考文献	(234)
第六章	事实和数值型数据库	(235)
第一节	事实和数值型数据库概述	(235)
1.1	事实和数值型数据库的发展历史	(235)
1.2	事实和数值型数据库的含义、特点与作用	(236)
1.3	事实和数值型数据库的主要类型与内容特征	(237)
第二节	英文事实和数值型数据库举要	(240)
2.1	DIALOG 商情数据库	(240)
2.2	人类生物基因组数据库	(240)
2.3	贝尔斯坦/盖墨林化学数据库(Beilstein/Gmelin CrossFire)	(242)
2.4	ISI 化学数据库(ISI Chemistry)	(246)
2.5	Gale 出版集团的参考性资料数据库	(250)
2.6	LEXIS-NEXIS 参考资料数据库	(254)
2.7	《不列颠百科全书》网络版	(255)
第三节	中文事实数值数据库举要	(257)
3.1	万方数据库资源系统的事实和数值数据库	(257)
3.2	中国科学院“科学数据库”	(258)
3.3	中国资讯行数据库	(259)
3.4	新华社多媒体信息数据库	(262)
3.5	“中国大百科全书”图文数据光盘数据库	(265)
第四节	其他事实与数值型信息资源	(267)
4.1	网上事实与数值型信息资源的评估	(267)
4.2	网上事实与数值型资源站点选介	(268)
	参考文献	(270)
第七章	电子期刊	(271)
第一节	电子期刊概述	(271)

1.1	电子期刊的特点	(271)
1.2	电子期刊的出版与服务	(272)
1.3	核心期刊和同行评审刊	(274)
1.4	电子期刊的检索	(274)
第二节	著名出版商的英文电子期刊	(277)
2.1	Elsevier Science 出版社的电子期刊	(277)
2.2	Kluwer Academic 出版社的电子期刊	(280)
2.3	Academic Press 的电子期刊	(283)
2.4	John Wiley 电子期刊	(284)
2.5	自然(Nature)	(286)
2.6	Springer 出版社的电子期刊	(288)
第三节	学会版英文电子期刊	(290)
3.1	科学在线(Science Online)	(290)
3.2	英国皇家物理学会(IOP)的电子期刊	(293)
3.3	英国皇家化学学会(RSC)的电子期刊	(293)
3.4	美国物理学会(APS)和美国物理研究所(AIP)的电子期刊	(294)
3.5	美国数学学会(AMS)的电子期刊	(294)
3.6	美国化学学会(ACS)的电子期刊	(295)
3.7	美国计算机学会(ACM)的电子期刊	(295)
3.8	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	(295)
3.9	美国心理学会(APA)的电子期刊	(295)
3.10	Jstor 英文过期期刊	(295)
3.11	缪斯计划(Project Muse)——以大学出版社为主的学术电子期刊	(297)
第四节	中文电子期刊	(297)
4.1	中国期刊网	(299)
4.2	万方数字资源系统电子期刊	(302)
4.3	维普电子期刊(中文科技期刊数据库(全文版))	(302)
4.4	台湾版电子期刊(光盘数据库)	(303)
第五节	网上免费电子期刊举要	(303)
5.1	斯坦福大学 HighWire 出版社的电子期刊	(303)
5.2	美国国家航空和宇航局电子期刊网站(NASA Astrophysics Data System)	(304)
	参考文献	(304)
第八章	电子图书和报纸	(305)
第一节	电子图书和报纸概述	(305)
1.1	电子图书	(305)
1.2	电子报纸	(310)
第二节	网络图书馆及其电子图书服务	(314)
2.1	美国“网络图书馆”(NetLibrary)电子图书	(314)
2.2	国内的中文电子图书系统	(315)
2.3	其他电子图书	(326)
第三节	电子报纸及其利用	(327)
3.1	重要电子报纸举例	(327)
3.2	其他电子报纸	(329)
	参考文献	(330)
第九章	网上其他学术资源的利用	(331)
第一节	搜索引擎的比较与应用	(331)

1.1	概述	(331)
1.2	常用英文搜索引擎及特点	(333)
1.3	常用中文搜索引擎及特点	(339)
1.4	搜索引擎的发展趋势	(343)
第二节	网络学术资源导航及其应用	(344)
2.1	概述	(344)
2.2	网络学术资源导航的建设过程	(346)
2.3	网络学术资源导航的内容	(347)
2.4	国内外网络学术资源导航系统介绍	(347)
2.5	网络学术资源导航的未来发展	(353)
2.6	如何在学术研究中利用网络资源导航	(354)
第三节	FTP资源及其使用	(354)
3.1	什么是FTP	(354)
3.2	FTP上的软件	(355)
3.3	使用FTP	(355)
3.4	FTP搜索引擎	(359)
3.5	常用的FTP站点	(361)
参考文献	(362)
第十章	数字信息资源的综合利用	(365)
第一节	课题查询及论文搜集资料	(365)
1.1	课题查询的方法	(365)
1.2	论文资料的搜集	(369)
1.3	检索示例	(370)
1.4	小结	(374)
第二节	学位论文开题及写作	(374)
2.1	学位论文开题及写作的特点、方法、步骤和格式	(374)
2.2	利用数字信息资源进行学位论文开题及写作	(377)
第三节	科技查新	(381)
3.1	查新的定义	(381)
3.2	科技查新的意义和作用	(382)
3.3	科技查新与一般课题查询的异同	(382)
3.4	科技查新与专家鉴定	(382)
3.5	查新程序	(383)
3.6	查新实例	(385)
参考文献	(388)
数字资源名称索引	(389)
数字资源分类索引	(393)

第一章 数字信息资源及其检索概述

引 言

计算机技术和远程通信技术的发展,使 Internet(互联网)在 20 世纪的最后十多年内飞速地蔓延到了全球 240 个国家。它通过采用统一的通信协议(TCP/IP),连接了世界范围内不同的信息系统,突破了地域、时空、文化、语言的限制,实现了跨国界、跨领域、实时的信息传递和交换。如今,网络 and 信息技术已经渗透到了文化、经济、政治、教育、科技、医疗、出版、新闻、体育、娱乐、商业以及社会生活各个领域,对社会、文化的发展产生了巨大的影响,人类认识世界、思考世界的观点、角度、方法都在不断地发生变化。

互联网的开放性及其信息资源共享和交换的能力,吸引了大量用户,越来越多的机构和个人在网上发布、查询和使用信息。据互联网协会(Internet Society,简称 ISOC)统计,1981 年,全球提供服务的主机(host)不到 300 个,而 2001 年 1 月,已达到 10957.4429 万个^①,用户达到 3 亿多人,已注册的域名 3600 多万个^②(其中注册为 .com 的商业机构约占三分之二),发展之快,远远超出了最初人们的设想和预测,其影响早已超出国界、文化、语言的界限。

在中国,互联网更是按几何级数的速度增长,据中国互联网络信息中心 2003 年 1 月公布的最新统计报告^③,截止到 2002 年 12 月 31 日,我国国际出口带宽数已达到总容量 9380 兆(MB),是 1997 年的 369.2 倍。上网计算机已经有 2083 万台,网上用户人数为 5910 万人,WWW 站点已有 37.16 万个,分别是 1997 年的 69.7 倍、95.3 倍和 248 倍。这些用户和网站的分布情况如下:中国公用计算机互联网(ChinaNET)、中国教育和科研计算机网(CERNET)、中国金桥信息网(ChinaGBN)、中国科技网(CSTNET)、中国联通互联网(UNINET)、中国网通公用互联网(CNCNET)、中国移动互联网(CMNET)、中国长城互联网(CGWNET)、中国卫星集团互联网(CSNET)。

互联网的发展,使信息环境发生了巨大的变化,信息从产生、传播、到使用及其种类和数量都与传统的资源大为不同。主要表现在:

1. 信息的产生

(1) 信息产生的速度加快,数量动态性持续增长。人们常说的“互联网年”为 3~4 个月,也就是说,互联网上每 3~4 个月增长的信息相当于传统方式的一年的信息量,这种发展将会逐步超过传统传媒工具报刊、广播、电视。

(2) 信息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,具有很大的任意性和自由性。

2. 信息的传播

(1) 信息的传播速度大大加快,尽管信息没有集中存放,而是分布在全球各地,但由于通

① <http://www.isc.org/ds/WWW-200101/index.html>. ② <http://www.domainstats.com/main.html>.

③ <http://www.cnnic.net/develst/2003-1/>.

过光纤网络传递,过去需要几天才能获取的信息现在几分钟内就可以得到。

(2) 信息传播方式发生变化,传统的信息传播方式是通过书刊、广播、电视向受众单向传播,而现在是通过光纤网进行,因而信息与受众的关系也发生了变化。信息与受众之间的关系从单向灌输转为平等交流,受众可以自由选择所需信息。

3. 信息结构

(1) 信息的内容发生很大变化,由于学术团体、政府机关、商业部门、个人、民间组织等任何组织或个人都可以在网上发布信息,因而对信息缺乏控制和管理,没有认证和审核,使得各种信息,包括学术信息、商业信息、个人信息甚至有害信息混在一起。

(2) 信息的形式从传统的印刷型图书期刊逐渐变为数字化信息,不再仅仅是视觉和静态形式,而是多媒体和动态的,需要功能强大的计算机软件系统来进行管理和使用。

(3) 信息资源的种类也不再仅仅是传统的正式出版物,而是电子期刊/图书、非正式出版物、灰色文献(半公开出版物)、数据库、软件、新闻组、电子公告板(BBS)、FTP 等各类资源共同构成网络环境下的信息资源。

4. 信息的使用

(1) 通过网络获取信息可以不受任何时间、地点的限制,可以在任何地方、任何时间随时随地从网上获取信息。

(2) 信息可以重复使用,可以较容易地进行二次、三次深度加工。

由于信息环境发生的上述变化,使得人们获取信息的途径发生了巨大变化,从网上直接获取信息的用户已逐渐增多,从图 1-1(引自参考文献 4)可以看出这种变化所在。

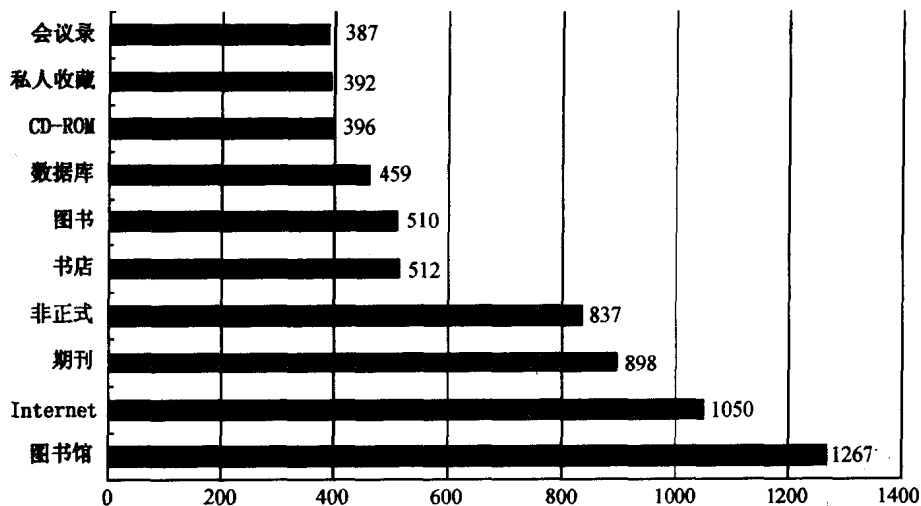


图 1-1 用户获取信息的主要途径

上述这些新的特点与变化都为人们获取和使用信息带来了新的问题:

(1) 信息的选择问题。海量信息从最初给人们带来的惊喜变成了一种令人无所适从、被淹没在信息海洋中的感觉,信息的无序和混杂,使用户不知道如何从网上快速准确地选择对自己有用的、高质量、正确的信息,不知道什么是查询信息的有效途径,有时花了很大精力和时间,却没有任何收获。

(2) 信息的检索问题。信息技术(IT)界和图书馆通过对信息的组织和管理,开发功能强

大而友好的检索系统,为用户提供了结构化的信息,以及有效的检索工具和检索途径。但用户是否具备检索能力,能否很快通过检索来获取和利用知识,仍然是目前多数用户面临的困难,因此对用户网络信息资源的检索和利用能力的培养就成为迫切需要解决的问题之一。

在此针对用户的具体需求再做进一步分析。中国互联网络信息中心 2001 年的调查^①表明,在网上用户中,按文化程度计算,本科以上文化程度者占 35.8%;按职业分析,专业技术人员占比例最大,为 20.6%;表 1-1 则说明用户的主要需求(仅列主要需求项目)。

表 1-1 网上用户需求分析 (%)

	新闻	计算机软硬件信息	休闲娱乐信息	电子书籍	科技教育信息
用户最希望获得哪方面信息	63.5	44.2	44.1	32.8	31.4
网上信息中哪些不能满足用户需求	19.3	19.9	18.3	29.1	21.1

从这些数字可以看出,上网用户中,学历较高者占三分之一强的比例;而用户对学术科研信息的需求量也很大,同样为三分之一左右,但在这方面的满足率并不高。这就说明,对于科研人员、专业技术人员、政府工作人员、教师、学生来说,更主要的问题是要解决学术和科研信息的查找和使用。

本书将针对上述需求,通过对数字信息资源,主要是学术资源的内容和使用的介绍,解决用户在网络环境下查找学术科研信息所面临的一系列问题,提高用户对信息资源的选择和检索能力,尤其在以下方面取得突破:

- (1) 阐述数字信息资源的结构与体系;
- (2) 各类学术资源的定义、特点及其应用;
- (3) 主要数据库、电子期刊、电子图书的学科范围、发展概况、特点及其具体检索方法;
- (4) 网络学术信息资源的综合应用。

第一节 数字信息资源的概念与类型

数字信息资源(digital information resources),狭义讲,亦可称为电子资源(electronic resources),指一切以数字形式生产和发行的信息资源。所谓数字形式,是以能被计算机识别的、不同序列的“0”和“1”构成的形式。数字资源中的信息,包括文字、图片、声音、动态图像等,都是以数字代码方式存储在磁带、磁盘、光盘等介质上,通过计算机输出设备和网络传送出去,最终显示在用户的计算机终端上。

随着互联网的发展,利用网络传递的数字信息资源的数量每年都以几何倍速增长,我们把这一类数字资源均称为网络信息资源(networked information resources),网络信息资源目前在数字信息资源中已经占有绝对比例。除此之外,到目前为止,仍然存在着大量仅在本地计算机上使用、没有通过网络传递的信息资源,如只用于单机的光盘或机读磁带数据库等,我们把这一类资源也归为数字信息资源。

数字信息资源不同于以往的印刷型文献资源和各类视听资料,与之相比,其特点主要有:

^① <http://www.cnnic.net/develop/2003-1/>.

(1) 存储介质和传播形式发生变化。数字资源可以将传统的图书、期刊中的文字、图片以及各类音像资料中的声音、动态图像融合在一起,利用数字技术进行制作,存储在光盘、磁带或硬盘等载体上。同时以网络作为主要的传播媒介,即转变为光信号,利用网络实现同步传输。不仅传播的速度大大提高,传递的信息量也超过了传统的出版物。例如一张光盘的最大存储量是 600 兆(MB),一套标准版的《不列颠百科全书》即可存储在一张光盘上;即使是多媒体版的也只需要两张光盘;而现在对数字信息的计算单位已经从“兆”变为“千兆”(GB)甚至“兆兆”(TB),一个数据库的容量通常是以 GB 或 TB 为单位计算的。正常速度下,从网上下载一篇几千字的文献最多只需要 1 分钟左右的时间。

(2) 以多媒体作为内容特征。集文本、图片、动态图像、声音、超链接等多种形式为一体,具体、生动、全方位地向用户展示主题,用户可以因此更加深入细致地了解所需信息的内容及其特征。

(3) 信息资源类型多种多样。既包括数据库、电子期刊、电子图书、电子报纸、专利等正式出版物,以及学位论文、教学课件等灰色文献,也涵盖了新闻组、电子公告板(BBS)等非正式出版的数字信息。信息交流的途径因此不再是单一化的,而是多层次、全方位的。

(4) 多层次的信息服务功能。数字信息资源最初产生时主要的服务功能是信息检索,发展到今天,已经产生了一系列的新功能:主动报道,如期刊目次报道服务;文件传递,如 FTP 服务;信息发现,如网络资源学科导航、分类主题指南等;网上讨论,如 BBS、新闻组等;这些服务功能扩展了传统出版物的职能,使数字信息资源得到更大程度、更深入的利用。

(5) 更新速度快、时效性强。传统的印刷型出版物一旦出版后,信息的内容就无法更改,必须要修订后出版新版本。而数字信息资源的更新和发布就容易得多,只要有人负责不断跟踪各个领域的最新发展变化,就可以随时修改内容,每月、每周、每日甚至每时更新,及时发布给用户。

(6) 具备检索系统,不再像传统文献那样需要逐页翻查,因而使用方便、快捷。特别是经过进一步加工的正式出版物,如电子期刊、数据库等,检索功能均很强大,可以很快找到自己所需的信息。

(7) 不受时间、地域限制,即没有收藏地点(如图书馆)、收藏时间(开放时间)的局限,可以随时随地存取。

1.1 数字信息资源的产生与发展

1.1.1 数字信息资源的产生与发展

数字资源伴随 1961 年美国化学文摘社(CAS)开始发行“化学题录”(Chemical Title)机读磁带而诞生。40 年以来,伴随着计算机和网络技术的发展,数字资源从无到有,从少到多,从书目型数据库发展到了全文数据库和电子图书、期刊、报纸乃至多媒体,从本地使用到网上发布,到今天最终成为人们生活和学习不可缺少的重要信息来源。

数字资源产生的最早的形式是数据库。1950 年代初,随着电子管计算机的产生,人们开始研究计算机情报检索系统,到 1958 年具有批处理能力的晶体管计算机产生后,计算机文献处理的研究开始有了突破性进展。1960 年代初,最早的数据库“化学题录”和“医学索引”(美国国家医学图书馆)相继产生。至 1965 年,据《Computer-Readable Databases: a directory and data sourcebook》一书统计,已有大约 20 个数据库可供使用,但这时的数据库存储介质仅限于机读磁带,内容以科

技文献书目、索引、文摘为主,检索也是以脱机批处理的方式进行,因此应用并不广泛。

1965年以后,由于集成电路计算机及硬盘的产生,以及数字通信技术和分组交换网的发展,开始有了数据库联机检索,著名的 DIALOG 系统以及 MEDLINE、ORBIT、BRS、JOIS 等相继开始服务,数据库的数量开始成倍增长,到 1975 年,已达到近 300 个数据库。数据库生产由政府行为逐步转向商业行为,用户也由原先的以政府机构为主扩展到更多的图书馆和科研机构,在内容上也开始增加人文社会科学和应用科学等内容。

20 世纪 70 年代以后,卫星通信技术、光纤通信技术、个人计算机的产生和发展给数据库联机检索创造了空前的发展机会,联机检索已不受地域限制,向国际化发展,个人用户开始加入到数据库检索行列中来。数据库的生产由美国向西欧扩展,在短短 10 年内即增长了 10 倍,到 80 年代末,数量已达到 3600 多个。数据库的容量增加,存储介质增加了光盘,因而也就产生了光盘数据库检索系统;数据库类型也有了变化,除以往的书目、文摘、索引数据库外,全文数据库开始迅速增加,而数值数据库、指南数据库等也已出现。

进入 20 世纪 90 年代,网络和信息处理技术的发展,使得基于互联网开发的数字资源及其检索系统有了突飞猛进的增长。仅从数据库即可看出这种发展之体现,以下为美国伊利诺依大学香槟分校图书馆情报学院教授 Martha E. Williams 对数据库发展的统计和归纳(数据截止到 1999 年):

(1) 数量: 参见表 1-2。

表 1-2 数据库增长情况

	1975 年	1999 年	增长倍数
数据库	301	11681	39
数据库生产者	200	3674	18
数据库代理商	105	2454	23
数据记录条数	5200 万	128.6 亿	242

(2) 类型: 打破了传统的文字、图片的单一类型,增加了集图像、声音、文字一体的多媒体数据库。其中文字型数据库中,在传统的书目、文摘索引、全文、事实数据库基础上,又增加了电子图书、电子期刊和报纸以及其他动态信息。全文数据库(含全文电子期刊、图书、报纸等)从 1985 年占全部数据库比例的 28% 增长为 1999 年的 50%,而书目型参考数据库从 1985 年的 57% 下降为 1999 年的 23%,主要原因是因为互联网和计算机技术的发展为全文数据库提供了传输的方便和海量存储的可能(参见图 1-2)。

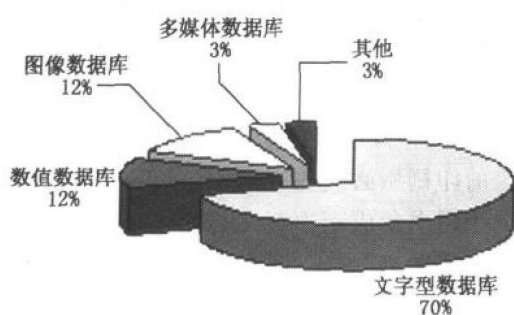


图 1-2 数据库类型分布

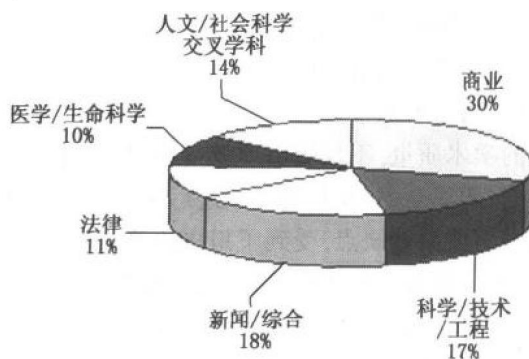


图 1-3 数据库内容分布