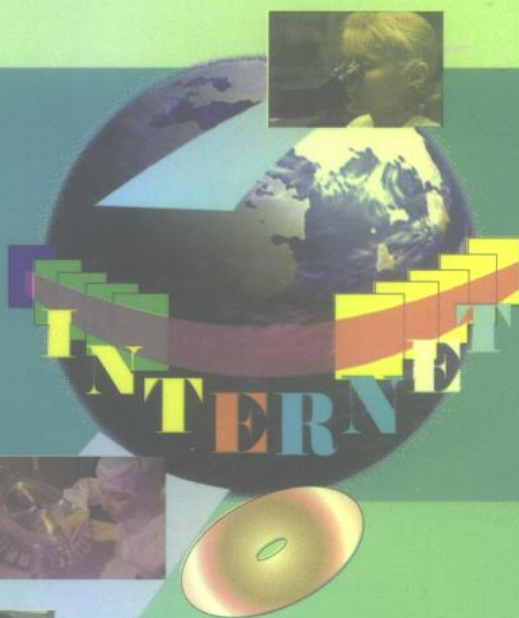


电子信息资源检索教程

主编 何翠华
编著 唐承秀 翟春红 张素芳

DIANZIXINXI
ZIYUANJIAN SUO
JIAOCHENG



南开大学出版社

TP393.4

428413

125

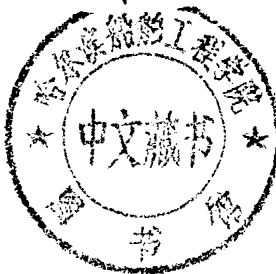
电子信息资源检索教程

主 编 何翠华
编 著 唐承秀 翟春红 张素芳



00428418

南开大学出版社



电子信息资源检索教程

- 主编** 何翠华
作者 唐承秀 翟春红 张素芳
责编 何志红
出版者 南开大学出版社
地址:天津市八里台南开大学校内
邮编:300071 电话:23508542
法律顾问 庞标
发行所 新华书店天津发行所发行
承印商 天津宝坻第二印刷厂印刷
版次 1998年11月第1版
印次 1998年11月第1次印刷
开本 787×1092 1/16
印张 19.125
字数 484千
印数 1-3000
ISBN 7-310-01175-9/TP·95
定价 25.00元
-

内 容 简 介

本书在简要介绍电子信息资源的历史与现状、类型与特点的基础上,讲述了 Internet 的起源与发展、入网方式和网上运行协议等知识,重点介绍了 Internet 上共享性信息资源系统和商业在线数据库的检索方法与步骤、应用策略与技巧,以及单机版、局域网和远程光盘数据库的检索方法与步骤。书中给出了许多上机检索实例,并按步骤给予指导说明。同时还介绍了 Web 主页的制作方法和 Internet 的三项基本服务功能: E-mail、Telnet 和 FTP。该书在附录中提供了常用 WWW 网址及其网页简介、Dialog 最新数据库目录及其常用数据库简介、OCLC 数据库简介和美国、加拿大部分大学入学办公室的 E-mail 地址。

本书既可作为高校师生“文检课”及“电子信息资源检索培训班”的教材,也可作为广大科技工作者、教师、学生和社会各界信息用户的参考书。

前言

电子信息资源检索是当代电子计算机技术、网络通信技术、信息产业和图书情报事业高度发展并且相互结合的产物,它于本世纪 60 年代发源于美国,现已发展成一种庞大而鼎盛的事业。可以说,衡量一个国家发展得快与慢、先进与落后的基本标准之一就是看其掌握、拥有电子信息资源检索的能力和程度,尤其是 Internet 的应用能力。

众所周知,Internet 是一个国际性的计算机网络。其中有不同学科、不同领域、不同方式、不同种类和不同性质的信息资料库。网上用户既可共享网中资源,也可将自己的信息资源发送到网上去。据《Internet Society》1997 年最新统计,世界上 238 个国家,目前已有 173 个加入了 Internet,个人用户数估计已超过 6000 万,并仍以每月 10% 的速度在递增。我国于 1994 年 5 月作为第 71 个成员国进入 Internet,虽起步较晚,但近一两年发展迅速。据中国 Internet 网络信息中心透露,截止到 1997 年 10 月 31 日,我国入网用户数为 62 万,而到 1998 年 5 月 31 日已达 106 万。Internet 正快速进入我国千家万户,对人们的工作、学习和生活产生了重大影响。但如何上网,如何使用网上的服务功能,如何查找和检索网上数据库中的信息,对大多数人来说还十分陌生。

在全球信息化的今天,电子出版物正挑战传统出版物,计算机检索正挑战传统手工检索,信息处理智能化、网络通信普及化和多媒体技术实用化已是社会发展的必然趋势。信息资源作为一种最重要的资源(与材料、能源相比)已逐渐为人们所认识,而电子信息资源又是信息资源的主体。因此,谁能及时地把握电子信息资源,谁就能够抓住发展的良机,从这个意义上来说,电子信息资源的检索能力又是人们在网络环境下生存和发展的必备素质。

目前所见的有关 Internet 网络的图书多是侧重于介绍其工作原理和基本服务功能的,没有详细地讲述如何检索网上信息资源(特别是商业在线信息资源系统),所以不能满足读者的信息资源检索需求。本书即是针对这些问题,把重点放在介绍各类电子信息资源的检索原理、策略、方法和步骤上。同时书中给出大量上机实例,并且每一步都给出指导说明,使读者根据其阐述,即可独立操作自己的计算机,或在 Internet 上漫游、检索,或在光盘数据库中浏览、查询。

本书曾作为多期“电子信息资源检索培训班”和近两年高校“文献检索与利用课”的教材,收到了很好的效果。但因水平有限,谬误在所难免,我们热切地希望广大读者不吝指教,以使本书日臻完善。

本书第 1 章由张素芳编写,第 2—5 章和附录 1 由何翠华编写,第 6—8 章和附录 2—4 由唐承秀编写,第 9—11 章和附录 5 由翟春红编写。全书由何翠华副研究员构思、编排和审定,唐承秀和翟春红二位老师参加部分章节的讨论和审定。

在编写过程中,我们得到李治安教授(博士生导师、南开大学图书馆馆长)、李文祯教授(南开大学网络中心主任、图书馆副馆长)和夏家善研究员(南开大学图书馆副馆长)的大力支持与帮助,在此谨表谢忱。

何翠华

1998 年 8 月于南开大学

目 录

第 1 章 电子信息资源概述	(1)
1.1 电子信息资源简介	(1)
1.1.1 什么是电子信息资源	(1)
1.1.2 电子信息资源的起源与发展	(2)
1.1.3 电子信息资源的类型和特点	(4)
1.1.4 我国电子信息资源建设的现状	(6)
1.2 电子信息资源利用的必要性及其培养途径	(8)
1.2.1 电子信息资源利用的必要性	(8)
1.2.2 电子信息资源利用的培养途径	(11)
第 2 章 Internet 概述	(14)
2.1 Internet 的起源与发展	(14)
2.1.1 什么是 Internet	(14)
2.1.2 Internet 的起源与发展	(16)
2.1.3 Internet 在中国	(17)
2.1.4 连接 Internet 的方法	(21)
2.2 Internet 运行方式与网络地址	(23)
2.2.1 Internet 运行方式	(23)
2.2.2 Internet 中的 TCP/IP	(24)
2.2.3 Internet 网络地址	(25)
2.3 Internet 的网络信息服务功能	(28)
2.3.1 Internet 的基本服务功能	(28)
2.3.2 Internet 的扩充服务功能	(29)
2.3.3 Internet 网上商业在线信息检索系统	(38)
第 3 章 WWW 信息系统资源搜索法	(39)
3.1 WWW 的服务功能	(39)
3.1.1 WWW 的编程语言	(39)
3.1.2 WWW 的服务功能	(40)
3.1.3 WWW 的网络地址	(42)
3.2 如何制作 Web 主页	(43)
3.2.1 HTML 文件的组成元素	(43)
3.2.2 如何建立 Web 主页	(48)
3.2.3 Web 主页源程序范例	(50)
3.3 WWW 系统资源搜索法	(54)
3.3.1 如何进入 WWW 系统	(54)
3.3.2 字符界面信息资源搜索法	(59)
3.3.3 图形界面信息资源搜索法	(61)
第 4 章 Archie 数据库检索系统的使用方法	(74)

4.1	如何进入 Archie 数据库检索系统	(74)
4.1.1	通过 Telnet 方式登录到 Archie 服务器上	(74)
4.1.2	通过 Archie 客户机程序访问 Archie 服务器	(76)
4.1.3	采用 WWW 客户程序访问 Archie 服务器	(77)
4.1.4	通过 E-mail 访问 Archie 服务器	(79)
4.2	Archie 系统命令与参数及其参数设置	(80)
4.2.1	Archie 系统的命令	(80)
4.2.2	Archie 参数及其参数设置	(84)
4.3	Archie 数据库系统的检索方法	(86)
4.3.1	字符界面下的 Archie 数据库资源检索方法	(86)
4.3.2	图形界面下的 Archie 数据库资源检索方法	(88)
第 5 章	Gopher 信息浏览系统的使用方法	(91)
5.1	如何进入 Gopher 信息系统	(91)
5.1.1	通过 Gopher 客户程序进入其服务器	(91)
5.1.2	通过 Telnet 进入 Gopher 服务器	(93)
5.1.3	通过 WWW 工具进入 Gopher 服务器	(94)
5.2	Gopher 系统命令	(95)
5.2.1	基本检索命令	(95)
5.2.2	文件管理命令	(96)
5.2.3	书签命令	(97)
5.3	Gopher 数据库系统检索方法	(97)
5.3.1	字符界面下的检索法	(97)
5.3.2	图形界面下的检索法	(105)
第 6 章	OCLC 信息检索系统的查询方法与步骤	(109)
6.1	关于 OCLC 系统	(109)
6.1.1	OCLC 系统简介	(109)
6.1.2	OCLC 系统的检索付费与计费方式	(109)
6.1.3	OCLC 系统的服务时间	(110)
6.2	OCLC 系统数据库的主题范畴及种类	(111)
6.2.1	主题范畴	(111)
6.2.2	数据库类型	(111)
6.3	OCLC 系统的检索方法与步骤	(112)
6.3.1	逻辑算符与位置算符	(112)
6.3.2	截词符号与禁用词	(112)
6.3.3	字段限定法	(113)
6.3.4	OCLC 系统样本库的使用	(113)
6.3.5	如何获取全文资料	(114)
6.4	OCLC 系统的联入方法与检索步骤	(115)
6.4.1	如何联入 OCLC 系统	(115)
6.4.2	WWW 环境下的检索方法与步骤	(115)
6.4.3	Telnet 环境下的检索方法与步骤	(119)
6.5	检索实例分析	(122)
第 7 章	Dialog 国际联机系统的检索方法与步骤	(125)
7.1	关于 Dialog 系统	(125)

7.1.1	Dialog 系统简介	(125)
7.1.2	Dialog 系统提供的服务方式	(125)
7.1.3	Dialog 系统的检索付费方式	(126)
7.2	Dialog 系统的联入方式与数据库种类	(127)
7.2.1	如何联入 Dialog 系统	(127)
7.2.2	Dialog 系统数据库的信息种类	(127)
7.3	Dialog 系统信息检索方法与步骤	(128)
7.3.1	逻辑算符与位置算符	(129)
7.3.2	字段检索法与截词检索法	(129)
7.3.3	Dialog 系统常用检索指令	(130)
7.3.4	WWW 环境下的 Dialog 系统信息检索	(131)
7.3.5	Telnet 环境下的 Dialog 系统联机检索	(134)
7.3.6	检索结果的处理及原文获取	(138)
第 8 章	光盘数据库检索	(140)
8.1	光盘简介	(140)
8.1.1	光盘数据库的由来与发展	(140)
8.1.2	光盘数据库的种类与特点	(140)
8.1.3	光盘数据库在我国图书情报工作中的应用	(142)
8.1.4	我国 CD-ROM 信息服务系统的几种模式	(142)
8.1.5	光盘数据库的一般检索步骤	(144)
8.2	常用中文光盘数据库检索方法与步骤	(144)
8.2.1	中文科技期刊数据库	(144)
8.2.2	中国化学化工文献数据库	(147)
8.2.3	人民大学复印报刊资料全文数据库及其索引光盘	(149)
8.2.4	中国学术期刊全文数据库	(153)
8.3	常用西文光盘数据库检索方法与步骤	(157)
8.3.1	美国 UMI 公司系列光盘数据库	(158)
8.3.2	美国银盘公司系列光盘数据库	(163)
8.3.3	Dialog 系统系列光盘数据库	(172)
第 9 章	E-mail——电子邮件	(179)
9.1	E-mail 简介	(179)
9.1.1	E-mail 的地址	(180)
9.1.2	E-mail 的结构	(180)
9.2	E-mail 的使用方法与步骤	(182)
9.2.1	在 DOS 下收发 E-mail 的方法与步骤	(182)
9.2.2	在 Windows 3.2 下收发 E-mail 的方法与步骤	(191)
9.2.3	在 Windows 95 下收发 E-mail 的方法与步骤	(195)
9.2.4	E-mail 的使用技巧	(203)
9.3	E-mail 的其他功能	(215)
9.3.1	用 E-mail 访问 FTP	(215)
9.3.2	用 E-mail 访问 Archie	(217)
9.3.3	用 E-mail 访问 Gopher	(219)
9.3.4	用 E-mail 访问 WWW	(220)
第 10 章	Telnet —— 远程登录	(222)

10.1 Telnet 简介	(222)
10.1.1 什么是 Telnet	(222)
10.1.2 Telnet 的功能	(223)
10.2 Telnet 的操作方法与步骤	(223)
10.2.1 字符界面下 Telnet 的操作方法与步骤	(223)
10.2.2 图形界面下 Telnet 的操作方法与步骤	(227)
10.2.3 Telnet 常用命令	(229)
第 11 章 FTP —— 文件传输	(232)
11.1 FTP 概述	(232)
11.1.1 什么是 FTP	(232)
11.1.2 FTP 的功能	(233)
11.1.3 如何连接 FTP	(233)
11.2 FTP 的操作方法与步骤	(235)
11.2.1 字符界面下 FTP 的操作方法与步骤	(235)
11.2.2 图形界面下 FTP 的操作方法与步骤	(244)
附录	(248)
1. 常用 WWW 地址及其简介	(248)
2. 常用 OCLC 系统数据库介绍	(268)
3. 常用 Dialog 系统数据库介绍	(274)
4. Dialog 系统所有数据库列表	(276)
5. 美国、加拿大部分大学入学办公室 E-mail 地址	(289)
参考文献	(297)

第1章 电子信息资源概述

电子信息资源自1946年世界上第一台大型电子计算机“埃尼阿克”(ENIAC)诞生之后得以从信息资源中脱颖而出,并在计算机技术、网络技术、通信技术和存储技术飞速发展的形势下得到了蓬勃发展。与其他各类信息资源相比,电子信息资源以其信息更新速度快、易检索、易利用和共享性好等特点对社会经济、科研、教育、生活、办公、生产和管理等方方面面正产生着巨大的影响。它日益成为各种信息资源中最重要、最活跃的一种,可以说,它是信息资源的主体,它将为人类带来新的财富、新的效率和新的生产力。如何熟悉、掌握其检索方法与策略,及时地获取和利用它为自己的工作、学习和生活服务,是当前我们所面临的重要课题。

1.1 电子信息资源简介

1.1.1 什么是电子信息资源

电子信息资源是信息资源的一种,是电子化了的信息资源,即以电子数据的形式,把文字、图形、图像、声音等多种形式的信息存放在光、磁等非印刷型介质上,以电信号、光信号的形式传输,并通过相应的计算机和其他外部设备再现出来的一种信息资源。

把信息视为资源是人类认识上的一大飞跃。以往人们只认为物质和能源是资源,现在人们则进一步认识到信息也是一种资源,而且是一种特殊形式的资源,是一种十分重要的资源,它与物质资源、能源资源一起并称为后工业社会即信息社会的三大支柱。与物质资源、能源资源相比,信息资源具有不可替代的长处。物质资源和许多能源都是非再生的,将随着人类的大规模使用而日渐枯竭,信息资源则随着人们的开发和利用而增加,是取之不尽、用之不竭的。目前,信息作为战略资源对经济、政治、科技、文化、军事等方面发展的重要性已得到世界各国的普遍认可,人们对资源的争夺已由占有土地、矿藏等自然资源逐渐转移到争夺信息资源。在我们大踏步地迈向信息社会的今天,信息资源重要性的观念正在深入人心。而信息社会的本义,就是利用现代电子信息基础设施快速、方便、准确地传输信息资源。这种能够利用现代电子信息基础设施来传输、记载的信息资源则是我们所说的电子信息资源。为了能清楚地理解电子信息资源的内涵,还是让我们循迹而择,沿着信息革命的轨迹来看一看。

自从有了人类,人们对信息的不自觉认识、开发及利用的历史就开始了。原始社会的“结绳记事”是人们存取信息的一种方式,在共同劳动和相互交往的过程中人们逐步创造了语言,语言成为信息交流的工具,信息从此得以口耳相传,可以说,语言能力的获得是人类历史上的第一次信息革命。

文字的出现使人类第一次突破时空的限制,获取、保存和流传了大量的信息,使口耳相传的信息固定下来,促成了信息的完整和大量积累,并能够世代相传。印刷术的发明则扩大了文字的作用。公元 11 世纪,中国的毕昇发明了活字印刷,取代了效率较低的雕版印刷。活字印刷术传到欧洲后得到了迅速发展。1438 年,约翰·古登堡(J. G. Gutenberg)发明了活字印刷机。之后,印刷术不断发展,经历了木印、石印、铅印、凹印、平印和胶印等不同的历史时期,使文字信息的大量复制成为可能,文字和印刷术的发明使文字成为信息的载体并使信息大量传播,这是人类历史上的第二次信息革命。

1844 年 5 月 24 日,电报的发明家莫尔斯发出了“上帝创造了何等的奇迹!”的电文,电报通信的时代开始了。这是人类第一次用电作为载体来传递信息,较之“烽火戏诸侯”时的烽火台,用“急脚递”传递信件的古代邮驿,信息传递速度大大加快。1866 年英国长达 4000 公里的大西洋海底电报电缆铺设成功,这是人类有史以来实现万里瞬间通信的开端。接着,电话、无线电、广播、电视、传真机和卫星通讯等通讯工具相继产生,这是人类历史上的第三次信息革命。

计算机的发明则引发了人类历史上的第四次信息革命。1946 年,美国莫尔电子工程学校和宾夕法尼亚大学的电子计算机设计组成功研制了世界上第一台大型电子计算机“埃尼阿克”(ENIAC),其设计思想来自数学家冯·诺依曼。在他的方案中,采用二进制的 0 和 1 来表示数据和指令,确立了数字电子计算机的体系结构。虽然计算机的发展十分迅速,半个世纪以来,其主要部件已经历了电子管—晶体管—集成电路—超大规模集成电路四个发展阶段,但冯·诺依曼的设计思想仍沿用至今。计算机的处理能力呈指数增长趋势,其应用也由最初的数据处理发展到信息处理,正是因为计算机的这种飞速发展和广泛应用,才使得信息资源的形式发生了变化,主要是信息资源在处理与存储时的形式发生了巨大的变化。信息资源以数字化的形式存在,也就是说,计算机实现了信息数字化的突破,所有的信息都可以用 1 和 0 这两个数字来表示,从而使信息的载体和传输介质产生了质的飞跃。计算机在加工处理信息的数量、质量、效率和速度等方面都大大超过了人脑,为了利用计算机的优点来处理信息,人们逐渐将各种信息资源数字化。之后,人们又将迅速发展的通讯技术、网络技术、存储技术、多媒体技术与计算机技术结合起来,现代信息技术开始兴起,网络化、数字化的信息环境加速形成,信息资源数字化成为必然。

由信息革命的发展历程我们可以看出,每次信息革命的发生都源于记载、传输信息载体的巨大变化。电子信息资源就是在这种变化之中产生的,它是现代信息技术发展的必然产物。现代信息技术将文字、图形、图像和声音等高密度信息数字化,以高速度、大容量传递到每一个办公室、教室、图书馆及家庭,使全世界的每个角落都能通过电子信息资源实现瞬间信息共享。

1.1.2 电子信息资源的起源和发展

1. 起源——机编文献目录数据库信息阶段

“埃尼阿克”(ENIAC)采用 0 和 1 来表示数据和指令,从而拉开了数据信息电子化的序幕。但最初计算机的功能也只是进行数据的计算,随着计算机技术的不断进步,计算机才逐渐开始用于信息的处理,而信息只有积累到一定的量才能称之为资源,所以我们说,最初计算机进行处理的并不是电子信息资源。电子信息资源的真正起源是美国国立医学图书馆在 60 年代中期用电子计算机出版的《MEDLARS》(Medical Literature Analysis and Retrived System)数据库,它被用于医学文献的检索。随后,美国其他一些机构也相继使用电子计算机编辑文摘期

刊,从而在客观上编制成了机器可读的书目文档,称之为文献目录型数据库。这种文献目录型数据库由于使用计算机来编辑、检索文献信息,因此我们可以将机编文献目录型数据库的研制视为电子信息资源形成的开端。

2. 国际联机检索信息阶段

随着科技发展的需要,这种目录型数据库的专业面越来越广,除自然科学、社会科学以外,还包括经济及人文科学等。另外,还出现了与目录型数据库类似的存取数值型信息的数值型数据库和全文记载事实的事实型数据库等。但这些数据库只能进行批量检索,即检索服务部门把许多用户的课题及其要求加以汇总,进行批量检索,然后把检索结果邮寄给检索用户。显然,其缺点是用户不能进行人机对话及时修改检索策略,更不能及时获得检索结果。通信网络的出现和发展则弥补了这一大缺陷,它促进了电子信息资源远距离的传输和交换。首先是银行界、建筑行业以及商业界为了得到准确及时的行情和金融方面的信息,率先使用了计算机处理技术,并进一步推动了公用通讯网络与计算机系统相结合的国际联机检索系统。国际联机检索就是用户使用终端设备,远距离地从国际联机检索中心迅速而准确地获取电子文献信息,使知识或信息得到广泛而有效的传播和利用,其实质就是数据库和通信的结合。从六七十年代起,许多国家还先后建立了专门从事计算机检索的机构,如美国的洛克希德公司(Lockheed)和系统发展公司(System Development Corporation),英国的目录检索服务处(Bibliographic Retrieval Service)和图书馆自动化情报服务处(British Library Automated Information Service)与Infoline公司以及欧洲的DIMDI(Deutsches Institut für Medizinische Dokumentation und Information)和ESA(European Space Agency)等,这些机构都建有大量的数据库联机检索系统,都向全世界联机用户提供电子信息服务。其中,著名的系统有Dialog、ORBIT、ESA-IRS、JOIS、BRS、QUESTEL、OCLC、STN等。以Dialog系统为例,美国洛克希德飞机与导弹公司的Dialog系统是目前美国和世界上最大的检索系统。Dialog系统于1966年正式开展文献检索,当时其数据库存储的文献只有20多万篇。随着时间的推移,其数据库的数量在逐年增加,数据库范围涉及到自然科学和社会科学的近20个学科,50多种语言。文献类型有:学位论文、会议录、科技报告、政府文件、专利标准、经济预测、公司行业名录、私人文档和统计数据等。由于联机服务和数据库产品迅猛发展,使电子信息资源引起了人们的普遍关注。但国际联机检索系统能够传输、检索的仍主要是电子文献信息资源,还停留在电子文献信息阶段。

3. 光盘信息阶段

利用国际联机检索系统检索到的电子文献信息具有较高的实用价值。但国际联机检索费用昂贵,一般用户难以负担,人们开始努力寻求一种低廉的存储、检索电子信息的方式,光盘存储技术则适应了这一需求。光盘是70年代出现的一种新型存储设备。近十年来,由于许多尖端技术领域所取得的新成就,特别是光学、光电子技术、微电子技术的发展,光盘技术也得到了飞速的发展。由于光盘在存储电子信息资源方面具有记录密度高、容量大、成本低、体积小、寿命长、可实现随机存取和检索费用低廉等优点,因此,光盘被广泛用于存储、检索电子信息资源,并产生了一批生产系列光盘的公司,如美国UMI公司和银盘公司等。光盘记载的电子信息资源并不局限于文献信息,还包括各种软件,但可用于检索的仍以文献信息为主。

4. 网络信息阶段

国际联机检索和光盘检索为我们提供了大量的电子信息资源,但各自又都有着或多或少的缺点,例如国际联机检索的费用昂贵,检索指令复杂,多为专业检索人员才会使用,而光盘检索得到的信息又不是十分及时等。因此,人们渴望一种新型检索方式的出现。1993年,美国政

府提出 NII(国家基础设施建设)计划,兴建以 Internet 为雏形的“信息高速公路”。随着 Internet 的迅猛发展,网络资源如潮水般涌来,它已成为人类进行科学研究、商业活动和共享信息资源的重要手段,并为世界各国所接受。Internet 是网络的网络,是网络通过互联而形成的全球网。它已延伸到地球上几乎每个国家。在 Internet 网上的所有主机都采用 TCP/IP 协议连接和通信,使网上各种计算机都遵循该协议所规定的方式进行数据交换,其结果是使得 Internet 有极为丰富的电子信息资源,堪称世界上最大的图书馆或信息资源库。Internet 信息资源主要包括电子报刊、电子新闻、电子报告、电子论坛、会议资料、各种软件资料、图像文件、声音文件和电子游戏等。由于 Internet 起源于美国军方,并在学术界得到了迅速发展,美国政府扶持的力度较大,因此网中大部分资源是免费的,如 Archie 数据库、WAIS 数据库、大多数大学图书馆的书刊数据库以及 WEB 信息系统中的共享性数据库等,可以说,使用 Internet 所提供的电子信息资源费用较为低廉。另外,Internet 的信息更新速度快,其信息在随时刷新,用户得到的信息总是最新的。此外,Internet 中的 WWW 网络信息系统采用链式结点链接,用户检索起来不但十分方便,而且检索范围大大拓宽了,可检索的信息也不再以文献信息为主。同时,Internet 还提供了多种信息服务,主要有 E-mail、Telnet、WWW 和 BBS 等。为便于充分利用丰富的信息资源,Internet 上还提供各种各样的检索工具,如 Archie 数据库检索系统和 Gopher 菜单式数据库检索系统等。总之,Internet 是目前世界上资料最多、门类最全、规模最大的信息库,是人们获取信息的重要来源。

1.1.3 电子信息资源的类型和特点

1. 电子信息资源的类型

电子数据分为以模拟信号存储的数据和以数字信号存储的数据,由此,电子信息资源也可分为两大类:模拟信息资源和数字信息资源。

模拟信息资源是通过广播、电影、电视和录像等以模拟信号来传递的信息资源。数字信息资源是通过电子计算机等以数字信号来传递的信息资源。数字信息资源是电子信息资源的主体。因此,我们目前所说的电子信息资源主要指的是数字信息资源,本书中讲到的电子信息资源检索也主要指的是对数字信息资源的检索。

数字信息资源按不同的标准可分为若干类型。

(1) 按信息的载体分

① 联机网络信息资源。本世纪 60 至 70 年代,世界上发达国家和地区相继建立起计算机联机信息服务系统,如美国的 Dialog、欧共体的 ESA 和德国的 STN 系统等,都为全世界联机用户提供了丰富的电子信息资源。但长期以来,由于费用昂贵,许多用户不敢问津。90 年代以来,Internet 迅猛发展,网络资源十分丰富,价格相对低廉,且有许多免费资源,这使得人们越来越多地从网络上检索、获取信息,通过网络来共享全球的信息资源。

② 单独发行的信息资源,以光盘出版物为主。光盘是一种用激光来记录和再现信息的高密度存储介质。1966 年,用光盘存储信息的概念首次被提出。1985 年,世界上第一个商品化的 CD-ROM 数据库——BIBLIOFILE(美国国会图书馆的 MARC 即机读目录)推出。光盘技术迅速发展,使其以容量大、体积小、重量轻、价格低并可同时存储声音、图像、文字,携带、邮寄方便等优势,使信息处理、传播、检索和利用的模式发生了革命性的变化。从信息存储的方式分类,光盘可分为模拟式光盘和数字式光盘两大类。从读写方式分类,光盘可分为 CR-ROM 只读式

光盘、WORM 一次写多次读光盘和 ORAM 可擦除式光盘三种。光盘版信息资源主要包括软件、游戏、娱乐、教育和各种文献信息数据库等。

(2) 按信息的表现形式分

① 电子出版物信息资源。电子出版物,指以电子方式或机读方式生产和发行的出版物。电子出版物的文字、图像和声音信息是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。目前出版的各种电子出版物有电子期刊、电子报纸、电子图书、电子名录、电子地图、数字声音(如 CD)和数字图像(如 VCD、DVD)等。除此之外,还有相当于二次文献的磁带式和以联机方式查询的数据库,如书目数据库、数值数据库、事实数据库和全文数据库等。

② 非电子出版物的电子信息资源。它主要指的是电子新闻、电子邮件、电子布告和电子论坛等。

(3) 按信息的媒体形式分

① 文本信息资源。普通的文本信息资源的知识单元按线性顺序排列。阅读时,人们跟随文本的线性流向吸收其中的养分,遇到不懂的地方或想要知道详细情况时,就得暂时中断阅读,去查阅有关参考资料,这就打乱了文本固有的线性配置格局,在读者的头脑中形成了相互参阅的知识单元网状结构。然而,用户不易掌握和追踪这种网状结构,更难以对其修改和补充,仅靠手动、眼看、心记是具有极大的局限性和片面性的。超文本的出现为解决这一问题提供了手段。

② 超文本信息资源。超文本是一种人-机交互的友好系统。用户利用计算机可以增删内容,用户的想法可随时存入数据库中,也可随时检索、调用。超文本是一种通信,它可以组织许多研究人员交流思想,沟通情况,这相当于开一个小型电子会议。一些有关的研究人员和设计人员,借助超文本可以共同撰写研究报告和设计工程项目。超文本信息资源是按知识单元及其关系建立的知识结构网络。其数据库由结点和链路组成,查阅超文本信息资源时,以知识片段及其关系作为追踪、检索的依据。除了处理一般的文字信息外,还包括图片、地图和其他直观信息,超文本信息资源能够把文字信息和图像信息有机地结合在一起。

③ 多媒体信息资源。多媒体是包括文本、图像和声音在内的各种信息表达或传播形式的总称。多媒体信息系统能针对用户的需求提供各种形式的信息。它们可以是文本、图像(图表、图画、照片、动画或活动影视)、声音(语言、音乐或其他音响)以及它们的结合。由于计算机软硬件技术的限制,相当长时间以来计算机信息检索系统只限于存储和检索书目、文摘等线索型文献,多媒体的出现使得人们接受的信息资源不但图、文、声并茂,而且丰富多采。

④ 超媒体信息资源。超媒体则是超文本与多媒体两种技术的结合。一般说来,当超文本节点中的信息是多媒体信息时,即在信息浏览环境下超文本的信息管理方式与多媒体的信息表现方法结合在一起时,就称为超媒体,它是超级媒体的简称。近几年来,超媒体技术发展迅速,在 Internet 上超媒体应用系统不断涌现。在超媒体信息系统中,不同类型的媒体信息能高度综合和集成,空间上图、文、声并茂,时间上媒体信息同步实现,有超文本和多媒体两种信息资源的特点,具有高度的交互性。

2. 电子信息资源的特点

(1) 信息存储形式为文本→超文本→多媒体→超媒体。这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之间的逻辑关系,为网络环境下不同形式的信息资源的管理和开发利用提供支持,由传统的顺序、线性排列,通过利用数字化存储技术,发展到超文本、超媒体技术,使得信息可按照自身的逻辑关系组成相互联系的、

直接的、非线性的网状结构。也就是说,用户不必关注所查信息存在何方,该项工作由 Gopher、WAIS、WWW 等检索工具自行去完成,为方便用户检索、提高电子信息资源的使用率奠定了基础。

(2) 存储介质发生转换。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号,从模拟信号转变为数字信号,使信息的存储传递和查询更加方便,且存储信息密度高,容量大,可以无损耗地被重复利用。

(3) 以现代信息技术为记录手段,是一种数字化的信息资源。信息以数字化的形式存在,既可在计算机内高速处理,又可借助通信网络进行远距离传播,这使得共享全球信息资源成为可能。而且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。

(4) 内容丰富。它既可是文字、图表等静态信息,也可是集图、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,把枯燥的文字信息转化为形式多样、活泼的数字信息,界面友好,易于人机沟通。

(5) 数据结构具有通用性、开放性和标准化的特点。它在网络环境下,可被数人同时访问,是一种共享性的信息资源。在讲求兼容性与标准化的信息社会中,数字信息资源易于实现信息资源的扩充,以及各信息系统之间的互联与互操作。

(6) 具有高度的整和性。电子信息资源不受时间、空间的限制,可以实现跨时空、跨行业的传播。

(7) 便于各种媒介信息的一体化。电子信息资源的相互转换和二次开发易于形成各种数据库,便于检索与使用,增大信息资源的利用价值。

(8) 交互式性能增强。由于数字信息资源存储在计算机能够识别的介质上,因此随着计算机软件的更新与性能的日益提高,用户逐渐具有更多的主动性。他们不仅是数字信息资源的利用者,而且将成为数字信息资源的开发主体。

1.1.4 我国电子信息资源建设的现状

1. 网络信息资源建设成就

网络信息资源建设的前提是网络信息基础设施的建设。自 1992 年美国政府的国家信息基础设施规划实施以来,全世界迅速形成了一个信息网络化建设浪潮。我国政府也及时作出反应,积极规划建设以“三金”工程为首的中国高速信息网。在政府的高度重视与支持下,经过“八五”期间的建设,中国信息化基础结构(CNII)有了长足的发展。这期间,新增干线近 10 万公里,微波干线 20 条,大型卫星地面站 19 座,一个以光缆为主体、以数字微波和卫星通信为辅助手段的大容量数字干线传输网正在形成。到 1996 年为止,由公用分组交换网(CHINAPAC)、数字数据网(CHINADDA)和公用计算机互联网(CHINANET)三大骨干网组成的公用数据通信网,已能覆盖全国 2000 多个城市,连通世界主要国际信息网络,共有各类用户端口 15 万个,用户达数 10 万。同时,各类专用信息也取得了相当的成果:中国教育科研网(CERNET)、中国经济信息网(CEINET)和中国科学技术信息网(CSTNET)的骨干工程建设已完成。各地区的通信网络也纷纷开通,如广东省信息网和海南省信息网、上海市交互网和天津市信息网等。另外,还有一批信息网络正在规划之中。网络信息基础设施的建设为信息资源的建设提供了有利的条件。

遗憾的是,我国的网络信息资源建设并不令人乐观。由于受“重硬轻软”观念的影响,网络信息资源建设没有提到应有的高度,“路上跑洋车”,“有路无车,有车无货,有货低质”的状况普遍存在。为了建设有中国特色的信息高速公路,开发我国自己的信息资源,1997年,由中国科技信息研究所牵头,会同全国各有关科技信息机构,应用市场机制共同建设一个具有综合性、权威性的资源网络——“中国信息(CIHNAINFO)”。其内容包括科技之窗、电子图书馆、经贸信息、政策法规、教育、人才市场、企业之窗、生活指南、读者之友和 Internet 集萃等。“中国信息”工程的近期任务主要包括:建成科技之窗和电子图书馆等重点栏目;从国内现有的 1000 多个数据库中选择有较高价值的数据库装入系统进行联机服务,可用通用环球网浏览器进行检索、查询;加速动态性的信息资源建设;组织各部委、省市信息机构合作建立“中国信息”,发挥联合的优势,利用三年时间使“中国信息”分中心遍布全国 30 个以上城市,形成规模和集团优势;尽快实现全国已有的 4000 多种科技期刊的上网服务。“中国信息”的远期目标是以世界著名的网络服务系统 American-Line 为样板,建立一个综合性、权威性的中国信息博览,用 5—8 年的时间发展成为具有数万用户的、中国最大的、信息资源最丰富的电子信息资源网络。我们期待着“中国信息”能够成为一个以科技信息为主体,集经济、金融、社会、文化、教育资源于一体,面向团体、机关、个人与家庭提供快捷、方便、全方位信息服务的综合性网络信息服务系统。

2. 光盘信息资源建设成就

目前的光盘出版物主要是 CD-ROM 光盘。CD-ROM 光盘作为一种不同于磁盘等磁性载体的光学存储器,由于它存储容量大、成本低和易用等特点而受到人们的普遍欢迎。它发展极其迅速,并成为多媒体、超媒体系统的主要存储介质;其数量成倍增长,仅 CD-ROM 光盘,1989 年全世界只有 200 种,1990 年增至 1025 种,1992 年为 3216 种,1994 年达 6700 种,到 1996 年则增加到 13000 种。随着光盘性能的不断提⾼及容量的不断增加,光盘出版物的增长速度继续上升,覆盖面越来越大,功能与品种也越来越齐全。从目前的发展趋势可以看出:今后越来越多的印刷型出版物将会被相应的光盘出版物所取代。

1992 年,《中文科技期刊数据库(光盘版)》出版发行,这是我国第一个采用光盘存储的中文数据库。其特点是用户界面友好,更新速度较快,为国家科研、教学和政府决策机构提供了一流的信息服⾮。1992 年 4 月 28 日,我国第一家开发制作多媒体光盘电子图书的专业公司——北京金盘电子有限公司诞生。目前,全国电子出版单位已发展到 41 家,具有一定规模的电子出版物开发制作企业达 100 多家,我国的光盘信息资源建设日益蓬勃发展。如数据库类的《中国法律法规检索系统》、《中国工商企业信息大全》、《中国企事业单位名录大全》、《计算机世界全文数据库》、《人民日报合订本》、《邓小平文选》和《中国百家报刊精选》等,教育类的《轻轻松松背单词》、《大嘴英语》和《CSC 电脑家庭老师系列》等,传统文化类的《中国古代美术》、《唐诗三百首》和《中华名胜》等,娱乐类的《赤壁》等游戏。值得一提的是,由光盘国家工程研究中心和北京清华信息系统工程公司联合组建的学术电子出版物编辑部主办、清华大学出版社出版的《中文学术期刊(光盘版)》,以现代信息技术手段为支撑,以 CD-ROM 光盘为载体,大规模集成我国各学科核心和有专业特色的现刊全文电子文本,配以具有国际先进水平的检索软件,构成了可以向国内外用户提供各种检索和专项咨询服务的大容量动态信息资源体系。这一期刊形态的资源型数据库,预计收录我国正式出版的期刊 3500 种,目前已收录 2600 种,且每月定期出版发行,每期文字量计 1 亿以上,分成 8 个专辑。该刊以“全文光盘检索管理系统”为支撑软件,具有全文检索、多窗口编辑、传真、打印、计时、自动收费等多项查阅管理功能,可在单机和局域网环境下运行。

3. 我国电子信息资源建设存在的问题

我国的电子信息资源建设虽然取得了一定的成就,但仍存在着许多问题,主要表现在以下几个方面:

(1) 电子信息资源建设严重滞后于信息基础设施建设。“重硬轻软”的观念是造成电子信息资源建设滞后的主要原因,从中央到地方各级政府部门对电子信息资源的发展缺乏必要的扶持。据有关部门统计,国家通讯设施到目前为止已投入 1000 亿元人民币,“七五”期间我国各部门投入信息系统建设的费用达 200 亿元,而用于电子信息资源开发(主要是数据库方面)的投资只有 5 亿—10 亿元,不到整个系统建设总费用的 5%。

(2) 电子信息资源的建设缺乏总体规划和必要的政策性扶持,处于无序状态。我国目前还未有一个权威性的机构对电子信息资源建设作出整体规划,有关电子信息资源建设发展的优惠政策也未出台,信息机构各自为政,重复建设、盲目建设,未形成规模,而且效益差。

(3) 电子信息资源建设结构不合理。我国的电子信息资源涵盖面窄,局限于行政、学术、科研领域,商用信息资源严重缺乏。以数据库为例,据《中国数据库大全》统计,1038 个数据库中,自然科学和工程数据库有 458 个,占总数的 44%,而经济类数据库只有 297 个,占总数的 29%。另据有关资料统计,目前全世界有商用数据库 5111 个,科技数据库只有 996 个,占总数的 19%,商业金融数据库占 50% 以上。由此可以看出,我国电子信息资源数据库多为科技型,社会急需的大型商业、物资、金融和经济数据库则十分缺乏,其建设结构不合理。

1.2 电子信息资源利用的必要性及其培养途径

1.2.1 电子信息资源利用的必要性

近年来,随着计算机、网络、通信与存储技术的迅猛发展,电子信息资源层出不穷,在整个信息资源中所占的比例越来越大,如何对电子信息资源加以利用,这个课题就被提到议事日程上来了。为了使每个社会成员都能自觉地、有意识地接受有关部门的电子信息资源利用的培训和自我培养,我们就有关电子信息资源利用的必要性问题从两个方面加以简单的阐释。

1. 我国社会发展的需要

现代社会最主要的特征就是信息化,信息化社会也称信息社会或后工业社会。由于信息革命使信息存储数量剧增,信息处理、传递速度大大加快,促进科学技术飞速发展,使社会各领域发生深刻变化。在信息社会中,计算机、通信和自动化技术组成一个完整的信息系统,把社会各部门联系起来,使人类的智力资源共享成为现实。最引人注目的是,信息社会中的一切社会活动似乎都被打上信息的烙印;信息成为经济发展和社会发展的最重要的战略资源;信息产业成为国民经济的主导和支柱产业,其产值和从业人数超过其他产业部门;价值的增长主要依靠信息的生产、收集和利用;信息消费程度明显增长等。可以说,信息作为一种最重要的资源,将成为一种能降低劳动强度及原材料、能源和财力等消耗的智力资源,是构成生产力最活跃的因素,也是科研、生产和管理人员进行创造性劳动的源泉。

随着世界新技术革命的兴起,信息时代的到来,人们逐渐认识到 21 世纪是信息化世纪,信