

数字信息资源检索

SHUZIXINXIZIYUANJIAN SUO

杨九龙 龚爱红 编著

陕西科学技术出版社

数字信息资源检索

杨九龙 龚爱红 编著

陕西科学技术出版社

图书在版编目 (CIP) 数据

数字信息资源检索/杨九龙, 龚爱红编著. —西安: 陕西科学技术出版社, 2008.8

ISBN 978-7-5369-4533-3

I.数… II.①杨…②龚… 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 (2008) 第 117880 号

内容简介

本书系统地介绍了数字信息资源的定义、类型、特点与历史, 数字信息资源检索的原理、方法与步骤, 搜索引擎的检索技术, 中外文常用数字信息资源的检索方法, 特种类型数字信息资源的检索, 网络资源的学科导航, 个人数字信息资源管理软件使用方法等方面的内容。可作为高等院校与科研机构的教师、学生、研究人员的培训教材和工作参考用书。

出 版 者 陕西科学技术出版社
西安北大街 131 号 邮编 710003
电话 (029) 87211894 传真 (029) 87218236
<http://www.snstp.com>

发 行 者 陕西科学技术出版社
电话 (029) 87212206 87260001

印 刷 西安长缨印刷厂

规 格 787 mm×1092mm 开本 16

印 张 20.25

字 数 42 千字

版 次 2008 年 9 月第 1 版
2008 年 9 月第 1 次印刷

定 价 40.00 元

版权所有 翻印必究

(如有印刷质量问题, 请与我社发行部联系调换)

作者简介

杨九龙 西北大学公共管理学院讲师，在读博士。主要著作有：《信息检索基础》（主编）、《数字图书馆原理与技术》（参编）、《信息描述》（参编）。

赵岩碧 西北工业大学图书馆研究馆员。主要著作有：《信息检索原理与方法教程》（主编）、《常用文献信息数据库检索指南》（第一著者）、《信息检索原理与方法教程》（主编）。

牛红亮 西北大学公共管理学院研究馆员。主要著作有：《中国图书馆年鉴》（参编）《陕西高校图书馆年鉴》（参编）《影响世界的一百个人》（参编）。

师俏梅 西北工业大学图书馆副研究馆员。主要著作有：《信息检索原理与方法教程》（参编）。

李桂影 西安交通大学图书馆馆员。主要著作有：《医学文献检索与利用》（副主编）、《现代医学文献检索与利用》（参编）、《医学信息检索与利用》（参编）。

龚爱红 西安财经学院图书馆馆员。

沈 思 西安科技大学图书馆馆员。

前 言

数字信息作为科学研究与知识创新的资源平台,其作用日益重要。如何高效全面获取所需的优质数字信息资源,已成为所有知识学习者与研究者迫切需要解决的问题。本书正是基于这种现实需求而组织一线文献检索课程教师集体编著的,旨在使信息需求用户能够在了解数字信息资源基本理论的基础上重点掌握核心数字信息资源的检索方法,更好满足其学习科研的需求。全书共分9章,由杨九龙与龚爱红编写大纲及最后统稿。其中第1章由杨九龙执笔;第2章和第3章由赵岩碧执笔;第4章的第1节至第4节、第8节至第9节、第11节至第12节由师俏梅执笔,第5节至第6节、第10节、第13节至第14节由杨九龙执笔,第7节由李桂影执笔;第5章的第1节至第5节、第7节、第15节由师俏梅执笔,第6节、第8节至14节、第16节由李桂影执笔;第6章由沈思执笔;第7章由牛红亮执笔;第8章第1节至第2节由师俏梅执笔,第3节由李桂影执笔,第4节与第6节由龚爱红执笔,第5节由沈思执笔;第9章由杨九龙执笔。

参加编著的作者是西北大学、西安交通大学、西北工业大学、西安科技大学、西安财经学院等西安地区知名高校长期从事文献检索课程教学与研究的富有经验的教师,本书反映了作者对数字信息资源检索长期的知识与经验积累,同时广泛吸收国内外数字信息资源检索方面的最新资料,力图能够最大程度的满足用户对数字信息资源检索理论与技术方面的知识需求。

由于时间仓促,加上作者水平有限与数字信息资源本身的迅速发展,不足之处在所难免,希望读者提出宝贵意见,以便今后修正完善。作者联系方式: yjlnwu@163.com。

感谢所有列出的和尚未列出的参考资源,这些高质量的研究成果给予作者以启迪。本书在编写过程中得到了陕西科学技术出版社李珑编辑的鼎力相助,在此特表深深的谢意和真诚的感谢。

杨九龙

2008年8月于西北大学新村

目 录

第1章 数字信息资源概述	(1)
1.1 数字信息资源的定义	(1)
1.2 数字信息资源的类型	(2)
1.3 数字信息资源的特点	(5)
1.4 数字信息资源的历史	(7)
1.5 信息资源与知识创新	(9)
第2章 数字信息资源检索基础	(11)
2.1 数字信息资源检索概念	(11)
2.2 数字信息资源检索语言	(11)
2.3 数字信息资源检索原理	(13)
2.4 数字信息资源检索系统	(15)
2.5 检索的方法途径及步骤	(19)
2.6 信息检索发展趋势	(27)
第3章 基于搜索引擎的信息检索	(31)
3.1 搜索引擎的概念与类型	(31)
3.2 搜索引擎一般检索方法	(34)
3.3 Google 简介及检索方法	(36)
3.4 百度简介及检索方法	(41)
3.5 常用中英文搜索引擎	(45)
第4章 中文数字信息资源检索	(57)
4.1 中文数字信息资源概述	(57)
4.2 中文数字书目资源检索	(58)
4.3 中文社会科学引文索引	(60)
4.4 中国科学引文数据库	(64)
4.5 全国报刊索引数据库	(66)
4.6 人大复印资料数据库	(69)
4.7 CNKI 系列数据库	(71)
4.8 维普知识资源系统	(78)
4.9 万方数据资源系统	(85)
4.10 中国资讯行数据库	(90)
4.11 中国高等教育文献保障体系	(92)
4.12 国家科技图书文献中心	(99)
4.13 中文数字图书资源检索	(104)
4.14 中文古籍数字资源检索	(112)

第 5 章	外文数字信息资源检索	(123)
5.1	外文数字信息资源概述	(123)
5.2	外文数字书目资源检索	(124)
5.3	Dialog 联机检索系统	(126)
5.4	STN 联机检索系统	(135)
5.5	CSA Illumina	(143)
5.6	ISI Web of Knowledge	(147)
5.7	OCLCFirstSearch	(154)
5.8	EBSCOhost	(158)
5.9	SciFinder Scholar	(163)
5.10	BIOSIS Preview	(168)
5.11	Free MEDLINE	(170)
5.12	Ei Compendex	(179)
5.13	Science Direct	(182)
5.14	INSPEC	(186)
5.15	IEEE/IEE Electronic Library (IEL)	(188)
5.16	Netlibrary	(192)
第 6 章	特种类型数字信息资源检索	(195)
6.1	标准文献及其检索	(195)
6.2	专利文献及其检索	(200)
6.3	学位论文及其检索	(210)
6.4	会议文献及其检索	(214)
6.5	科技报告及其检索	(217)
6.6	参考工具书及检索	(221)
第 7 章	人文社科数字信息资源导航	(236)
7.1	政治类数字信息资源导航	(236)
7.2	法律类数字信息资源导航	(240)
7.3	经济类数字信息资源导航	(245)
7.4	语言文学类数字资源导航	(252)
7.5	艺术类数字信息资源导航	(256)
7.6	历史类数字信息资源导航	(259)
第 8 章	科学技术数字信息资源导航	(262)
8.1	数学物理类数字信息资源导航	(262)
8.2	化学化工类数字信息资源导航	(270)
8.3	生物医学类数字信息资源导航	(275)
8.4	农业类数字信息资源导航	(280)
8.5	工程技术类数字信息资源导航	(286)
8.6	地球环境科学类信息资源导航	(294)

第 9 章 个人数字信息资源管理介绍 (297)

9.1 个人文献信息管理软件概述 (297)

9.2 个人数字图书馆概念及特点 (297)

9.3 个人文献信息管理软件使用 (298)

9.4 个人数字图书馆软件的使用 (310)

参考文献 (315)

第1章 数字信息资源概述

信息资源是信息资源管理的对象，是附加了人类劳动的富有价值的信息，是人类生存和社会发展不可缺少的资源。信息资源的开发利用程度是衡量一个国家信息化水平的重要标志。随着计算机技术、网络技术、通信技术的迅速发展和社会信息化程度的不断提高，信息资源的生产、存储和传递的方式发生了革命性的变化，传统的以手工为主的信息资源逐步过渡到现代的以数字化为主的信息资源。数字信息资源以其独特的优势使得人们对信息资源的经济价值和社会价值的认识不断提升，有效检索与利用数字信息资源的能力已成为科学研究、知识创新的基本素养。

1.1 数字信息资源的定义

数字信息资源也可称为数字化信息资源，其内涵有广义和狭义之分。广义的数字信息资源是指数字信息活动中所有要素的总和。具体说数字信息资源是以数字信息为核心，以信息人员作为收集、开发、传播信息的主体，以信息技术和信息设施作为获取、传播、开发、利用数字信息的方法、手段、工具和条件，是数字信息、人员、技术与设备的有机集合体。狭义的数字信息资源只限于数字信息本身，而不包括其他因素，应是有用的数字信息的集合，是经过人类筛选、组织、加工而形成的有序、可存取和能够满足人类需求的各种数字信息的集合。或者说狭义的数字信息资源是指以数字化形式记录的，以图像、文字、声音、视频等多种媒体形式表达的，存储在磁带、磁盘、光盘或分布式存储在因特网不同主机上的，可通过计算机本地检索、阅读或远程读取、使用的信息资源的集合。从信息检索与利用的角度出发，狭义的理解更符合用户的需求。根据数字信息资源所依附的载体看，包括电子信息资源（Electronic information resources）与网络信息资源（Network information resources）。电子信息资源是指存放于信息服务机构的数字化资源，有光盘、磁盘、磁带等各种载体形态，是近程存取的现实电子资源。而网络信息资源是指通过计算机系统及通讯网络才能获取的数字化资源，是远程访问的虚拟电子资源。

1.1.1 电子信息资源

电子信息资源的前身被称为计算机可读型文献或机读型文献，现特指电子化了的信息资源。可从以下三个方面来理解电子信息资源：首先电子信息资源是以磁性材料或其他材料为存储介质。这类载体具体有形，如软盘、磁带、硬盘、光盘、U盘等，其中软盘随着信息存储技术的发展已渐被淘汰，磁带通常主要用来备份不常用的数据以及用于在系统间传输数据，而硬盘与光盘则是电子信息资源的主要存储载体。其次以磁、光、电等高新技术

为记录手段,以电子形式存储在磁带、磁盘或光盘等载体上。存储在磁性介质上的信息是数字化的形式,模拟形式的信息需通过编码和程序设计输入计算机,使文字和图像信息转换为二进制代码的机器语言,存储在磁带、磁盘或光盘等载体上,阅读时再由计算机将其转换为文字或图像输出,显示在终端屏幕上。第三,存储电子信息资源的载体须与计算机等信息设备连接。利用特定的程序与软件方可检索与阅读其中的信息内容,具有信息的非人工识读性。同时电子信息资源相比传统的印刷型文献而言,信息存取的密度和效率更高、信息的可操作性与易检性更强。

概括地说,电子信息资源是以电子数据的格式把文字、图形、图像、声音等多种形式的信息存放在光、磁等非印刷型、非网络化的介质上,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。电子信息资源如按电子信息存储介质来划分,可分为磁心、磁鼓、磁带、磁盘、光盘等类型;如按电子出版物的类型划分,可分为电子期刊、电子图书、电子报纸、电子图谱或照片、光盘数据库电子产品、联机数据库电子产品。电子信息资源相比网络信息资源具有内容知识性、组织有序性、保存长期性等优点,是数字信息资源体系的重要组成。

1.1.2 网络信息资源

网络信息资源通常可理解为在计算机网络上以各种方式如博客、聊天、电子公告版、电子邮件、数据库、网站等交流和利用的所有静态与动态海量信息资源的集合,这类资源以数字化形式记录,以多媒体形式表达,存储在网络计算机磁介质、光介质以及各类网络通信介质上,并通过计算机网络通信方式产生、发布、存取和传递的海量存储型信息。与电子信息资源相比,网络信息资源是指网络系统范围内的信息对象,即在计算机网络系统中以各种方式产生、存在、发布并传播的信息内容的集合,此处的网络虽然包括局域网、城域网,但主要还是因特网,因为只有因特网上的信息资源才是最能够被广泛存取与利用的资源。

网络信息资源和传统信息资源的区别,表现为数字和非数字信息资源的区别、存贮方式分布与集中的区别、信息资源拥有和信息资源存取的区别以及交互型信息资源和非交互型信息资源的区别。网络信息资源已大大扩展了信息资源的原有内涵和外延,具有数量大、种类多样、管理的技术手段不断创新、网络信息资源传递快捷、时效性强等优点,同时也具有信息分散、无序,冗余信息、垃圾信息大量存在,信息结构复杂、安全性差等不足。

1.2 数字信息资源的类型

数字信息资源依据不同的标准可划分为不同的类型,对数字信息资源的分类以网络信息资源为主兼顾电子信息资源。

1.2.1 按信息加工层次分

1. 一次信息资源

是指人们直接以自己的生产、科研、社会活动等实践经验为依据生产出来的信息资源,也常被称为原始信息资源,其所记载的知识、信息比较新颖、具体、详尽。包括电子图书、电子期刊、电子报纸等,博客中的原创博文也属于一次信息资源。

2. 二次信息资源

是对一次信息资源进行加工、提炼和浓缩之后所得到的产物,是为便于管理和利用一次信息资源而编辑、出版和累积起来的工具性信息资源。包括搜索引擎、检索数据库、网络导航、学科指南等。

3. 三次信息资源

指对一次信息资源进行综合、分析、提炼后,形成的知识性、综述性信息资源。包括网络述评、参考数据库、网络推荐、网上百科全书等。

1.2.2 按信息表现形式分

1. 文本信息资源

文本信息资源是按知识单元的线形顺序排列,线性组织文本的方式是由人类思维的逻辑性和传统文本相对严密的结构决定的。但由于人类的阅读并非总是按单一顺序进行的静态过程,阅读更多的是跳跃式、不连续的、需要查询相关资料来理解正文的动态过程,而文本固有的线性配置格局无法满足读者的这种需求,难以在读者的头脑中形成相互参阅的知识单元网状结构。超文本的出现则有效解决了文本信息资源组织方式上的缺陷。文本信息资源如扫描版的古籍、WORD 文档、PDF 文件等。

2. 超文本信息资源

超文本是一种按信息之间关系非线性联想式存贮、组织、管理和浏览信息的计算机技术,它将文本信息存储在许多节点上,用链接将这些节点连成网状结构,节点是知识单元、片段,链接则表示知识之间的关系,因此超文本信息资源是按知识单元及关系建立的知识结构网络。检索超文本信息资源时,可从任何一个节点开始,以知识片段及其关系作为检索、追踪的依据。如网上的百科全书、字典等。但超文本信息资源在浏览时存在迷路、信息过载等问题,降低了资源的可用性。

3. 多媒体信息资源

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)以及它们的结合。多媒体信息资源在展示信息、交流思想和抒发情感等方面具有更好的效果。

4. 超媒体信息资源

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。超媒体除

了使用文本外, 还使用图形、图像、声音、动画或影视片断等多媒体来表示信息, 建立的链接关系是文本、图形、图像、声音、动画或影视片断等多媒体之间的链接关系, 实现了不同类型媒体信息的高度综合和集成, 空间上图、文、声并茂, 时间上媒体信息同步实现。

1.2.3 按人类信息交流方式分

1. 非正式出版信息资源

指流动性、随意性较强、信息量大、信息质量难以保证和控制的动态性数字信息资源。如电子会议、电子新闻、电子论坛、电子邮件、新闻组、电子公告版、聊天等网络活动所产生的各种信息, 但许多最新的最前沿的信息可能源于此。

2. 半正式出版信息资源

也叫灰色信息, 是指受到一定产权保护但没有纳入正式出版信息系统中的信息, 如政府机构和非政府组织提供的信息、学术团体和研究机构提供的信息、企业和商业的产品介绍、各种内部电子期刊、会议文集和各类报告等。

3. 正式出版信息

指受一定知识产权保护, 信息质量可靠, 利用率较高的知识性、分析性信息。这类信息多数为有偿使用, 如网络数据库、电子图书、电子报刊等。

1.2.4 按信息传输协议分

1. WWW 信息资源

WWW (World Wide Web, 简称 www 或 Web) 信息资源是建立在超文本、超媒体技术以及超文本传输协议 HTTP (Hyper Text Transfer Protocol) 的基础上, 集文本、图形、图像、声音为一体, 并以直观的图形用户界面展现和提供信息的网络资源形式。

2. Telnet 信息资源

Telnet 信息资源是指借助远程登录, 在网络通信协议 (Telecommunication Network Protocol) 的支持下, 可以访问共享的远程计算机中的资源。Telnet 使用户可以在本地计算机上注册到远程计算机中的资源。使用 Telnet, 用户可以与全世界许多信息中心、图书馆及其他信息资源联系。

3. FTP 信息资源

FTP 信息资源是指利用文件传输协议 FTP (File Transfer Protocol) 可以获取的信息资源。FTP 使用户可以在本地计算机和远程计算机之间发送和接收文件, FTP 不仅允许从远程计算机上获取、下载文件 (Download), 也可以将文件从本地计算机拷贝传输到远程计算机 (Upload)。FTP 是目前 Internet 上获取免费软件和共享软件资源不可缺少的工具。

4. Usenet/Newsgroup 信息资源

互联网上各种各样的用户通信或服务组是比较受欢迎的信息交流形式, 包括: 新闻组 (Usenet News Group)、邮件列表 (Mailinglist)、专题讨论组 (Discussion Group)、兴趣组 (Interest Group) 等。这些讨论组都是由一组对某一特定主题有共同兴趣的网络用户组成的

电子论坛，在电子论坛中所传递与交流的信息就构成了互联网上较流行的一种信息资源。

5. Gopher 信息资源

Gopher 是一种基于菜单的网络服务，它为用户提供了丰富的信息，并允许用户以一种简单的、一致的方法快速找到并访问所需的网络资源。全部操作是在一级级菜单的指引下，用户只需在菜单中选择项目和浏览相关内容，就可完成对 Internet 上远程联机信息系统的访问，无需知道信息的存放位置和掌握有关的操作命令。

6. WAIS 信息资源

WAIS (Wide Area Information Service) 是一种基于 Z39.50 标准的数据库索引查询服务或者说是一种分布式文本搜索系统，为用户提供了通过给定索引关键词查询到所需文本信息的服务，如文章或图书等。

1.2.5 按用户信息存取的方式分

1. 检索型信息资源

是指用于检索网络信息资源的检索工具，包括联机在线书目检索和搜索引擎等。

2. 邮件型信息资源

是指信息服务机构根据用户的需求采用电子邮件和电子邮件群体服务(Mailinglist)的方式进行传递的信息资源。

3. 揭示板型信息资源

是以不特定的大多数网络用户为对象的非即时的信息传递方式所产生的信息资源，如网络新闻和匿名文件传输 (FTP) 等。

4. 广播型信息资源

指在网络上向特定的利用者即时提供图像和声音信息的传播方式所产生的信息资源。

5. 电话型信息资源

指以某些特定的个人和群体为对象即时传递的信息，如网络参考咨询、会话 (Talk) 等。

除上述划分方法外，还可按信息资源提供的机构划分为图书馆、专业信息服务机构、企业公司团体及个人提供的信息资源。按所储贮的载体为标准分为磁盘型、光盘型、网络型信息资源。按信息发布形式可分为电子图书、电子期刊、电子报纸、数据库型、多媒体型等。按稳定性程度分为动态信息资源和静态信息资源。按内容涉及的领域可分为政府信息、学术信息、教育信息、文化信息、娱乐信息等。

1.3 数字信息资源的特点

1.3.1 信息格式数字化和存储载体大容量化

信息由存储于纸张上的文字符号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，这种数字化的信息以现代信息技术作为记录手段，使信息的存

储、传递、加工和查询更加方便,可以被无损耗重复利用,可在计算机内高速处理,又可借助于通信网络进行远距离传送。信息载体具有存储密度高、存储量大等特点,如 CD 盘可存贮 700MB 的数据, DVD 光盘的最大容量可达 17GB,单个硬盘的容量也有 400GB,而磁盘阵列的容量更大,这种大容量的存储载体满足了海量数字信息资源存贮的需求。如果采用网络附加存贮 NAS (Network Attached Storage) 和存储区域网络 SAN (Storage Area Network) 等技术,从理论上讲可满足无限数量数字信息资源存贮的需求。

1.3.2 信息组织网状化和信息结构复杂化

利用超文本与数字技术使得信息的组织方式发生了质的变化,即知识信息单元由传统的顺序线形排列转化为可按照自身的逻辑关系组成直接的、相互联系的、非线形的网状结构,知识信息单元间的网状结构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,这种组织方式使用户不必关注所查信息存贮位置,只须由搜索引擎、网络指南、学科导航等检索工具完成,便于检索与利用,提高了数字信息资源的利用价值。这种立体网状的文献链,把不同国家、不同地区、各种服务器、各种网页及各种不同的信息资源通过节点链接起来,大大增加了它的关联度。但由于网络信息资源广泛分布在世界各国不同地区的服务器上,而服务器又采用不同的操作系统、数据结构、字符集、图形界面、处理方式等,加之网络信息自身缺乏统一的标准和规范,使的信息结构极为复杂。这种复杂的结构增加了用户检索成本,且检索噪音难以降低,因此需要对目前主流的信息检索工具性能进行改进或开发高效的信息检索工具。

1.3.3 信息类型多元化和信息传播网络化

数字信息资源包括不同学科、不同领域、不同地域、不同语言,各种层次的、正式的和非正式的、动态的和静态的、数字化的和虚拟的各种信息模式和类型,并呈现出包括各种文本、表格、图像、声音、程序软件、数据库等多样性的信息形式,是一种多媒体、多语种、多形式、多类型、非规范、跨时间、跨地区、跨行业的混合体型信息资源。各类型信息共同并存,互为补充,形成整体,呈现出多元化的特征。同时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。数字信息资源不受时间、空间的限制,可以实现跨时空、跨行业的传播。随着网络化的普及和发展,网络信息资源的传播范围将会越来越广,人际交流的范围亦将不断扩大。

1.3.4 信息存取高效率和信息高度共享化

电子信息资源具有很高的信息存取速度。由于省去了传统的出版环节,并可利用数字技术随时对内容进行增、删、改等操作,故出版周期短、易于更新。数字信息的存取、传输速度快,且实现超文本链接,用户可方便快捷的检索、获取或阅读所需信息。用户利用

查询软件或检索工具,并可经多条检索途径随意地浏览或检索到所需信息,还可直接获取、阅读和交换相关数据库中的数据、图形、图像等信息,形成多层次、多角度的立体化查询。通过网络资源不仅可获取各种信息的线索,也可检索到有关信息的全文或原稿。同时还可获取大量的灰色文献或边缘文献信息,如调查采访报告、研讨会发言、各种原始笔记、项目计划报告或政策方针等信息。网络环境下,信息资源可被数人并发访问且无利用次数限制,实现了资源的高度共享,这种共享成本低、效益高,属无损消耗。

1.3.5 信息的可操作性与信息的交互性强

数字信息具有易变性,可以随意地进行操作,方便地改变其存在状态。如可根据需要得心应手地增删、修改文献信息的内容,或重新组织文献信息的结构;可调用存在于不同地方的数据组织动态文献;可进行信息压缩,从而提高存储密度或传输速度。可转换信息的媒体形式,如将数字文本经扫描转换成数字图像,或经光学字符识别后又转换成数字文本等。传统载体信息与用户之间是一种单向的关系,用户处于被动接受的地位,而网络信息资源具有互动性,用户的主动性增强,他们既是网络信息资源的利用者,也是网络信息资源开发的主体。用户可以利用网络信息高速公路、电子信箱、BBS 公告、各种论坛及帖子等交流手段发表自己的意见,进行各种形式的信息交流,真正在网上实现网络信息的资源共享。

1.3.6 信息质量良莠不齐且信息污染严重

由于网络信息增长迅速,导致信息新陈代谢加快、过剩甚至贬值,使得大量网上信息成为垃圾信息。网络共享性与开放性使得人人都可以在互联网上索取和存放信息,信息发布随意性强,信息提供者成分复杂,没有质量控制和管理机制,这些信息没有经过严格编辑和整理,良莠不齐,各种不良和无用的信息大量充斥在网络上,形成了一个纷繁复杂的信息世界,给用户选择、利用网络信息带来了障碍。他们在网上的活动又没有任何约束,且网络系统自身又缺乏必要的过滤、质量监控及统一完善的管理机制,这势必导致网络信息内容庞杂,信息质量良莠不齐,信息传播难以控制,黑客攻击频繁,计算机病毒肆虐,伪劣信息、无用信息泛滥成灾,信息污染变得越来越严重。

1.4 数字信息资源的历史

1.4.1 文献目录数据库信息阶段

早期计算机的功能集中在数据计算方面,随着计算机技术的不断进步,计算机逐渐开始用于信息的处理。20 世纪 60 年代初,各文摘社为克服因信息巨增而带来的困难,纷纷引进了先进的计算机技术。将经过整理、加工的文献信息输入到计算机中,由计算机进行编

辑和排版,输出后排版印刷为文摘刊物和各种索引。同时,仍保留在计算机中的机器可读的文献信息,作为二次文献编辑出版的副产品,发展成为文献目录数据库。由于机器可读信息大多记录在计算机的磁带上,因此也称文献目录数据库为磁带版二次文献。电子信息资源起源于美国国立医学图书馆(NLM)在1964年用电子计算机出版的《MEDLARS》(Medical Literature Analysis and Retrieval System)医学文献分析与检索系统,这是第一个大型的批式检索数据库。此后有美国化学文摘社的《化学文摘数据库》(CACON)等。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。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。

1.4.2 国际联机检索信息阶段

20世纪70年代以来,公用通讯网络与计算机技术、数据库技术相结合的国际联机检索系统进入实用阶段,实现了电子信息资源的远距离传输,使知识信息在世界范围得到广泛的传播和高度的共享。用户使用检索终端设备,通过通信设施,直接与中央计算机相连,检索远程数据库中的信息资源,其实质就是数据库和通信的结合。联机系统为全世界用户提供了丰富的电子信息资源,如著名的Dialog、STN和ORBIT系统,其中Dialog已成为世界上最大和历史最悠久的联机数据库信息服务系统,专业范围涉及新闻与媒体、商业与金融、知识产权、政府和法规、科学技术、能源和环境、医学、制药、化学、食品与农业、社会科学、电子信息行业等几乎所有的专业。

利用国际联机检索系统数据库得到的电子信息资源具有实时、快速和信息追溯年代长、查准率高,实用价值较高的优点。通过检索过程的人-机对话方式,可随机或脱机浏览、传递所得信息,并可根据需求随机修改检索策略。但联机费用昂贵,一般信息用户难以承受。光盘存储技术的出现较好的解决了国际联机检索费用昂贵的问题。

1.4.3 光盘信息阶段

20世纪70年代发展起来的光盘存储技术,具有存储密度高、容量大、重量轻、成本低、体积小、寿命长、可实现随机存取和检索费用低廉、携带邮寄方便等优点,同时可存储声音、图像、文字等多种类型的信息,使信息处理、传播、检索和利用的模式发生了革命性的变化。光盘是一种用激光来记录和再现信息的高密度存储介质。按信息存储方式可分为模拟式光盘和数字式光盘两类,按读写方式可分为CD-ROM(只读式光盘)、CD-R(只写一次式光盘)和CD-RW(可擦写式光盘)三种。用光盘作为信息存储载体的光盘数据库降低了电子信息资源的利用成本,得到广泛的应用。并产生了多家生产系列光盘的公司,如美国UMI公司和银盘公司等。

1966年,用光盘存储信息的概念首次被提出。1985年推出了世界上首个商品化的只读

式光盘数据库——BIBLIOFILE（美国国会图书馆的 MARC 即机读目录）。光盘信息资源主要包括各种信息数据库，它有单机版可进行单机检索，有网络版可在网上进行检索，还可以与联机检索系统联网进行联机检索。光盘信息资源的检索过程与联机检索极为相似，主要不同之处是信息追溯年代不如联机系统检索年代长，但检索费用大大低于联机费用，在通信不发达的地区，利用单机光盘信息资源是一种获取信息的有效途径。

1.4.4 网络信息阶段

国际联机检索和光盘检索在为用户提供丰富电子信息资源的同时，由于自身的局限并不能很好满足用户的需求，如国际联机检索的昂贵费用、复杂的检索指令和光盘信息资源的滞后性。Internet 以其信息资源数量庞大、增长迅速、更新速度快、检索便捷、免费信息丰富且网络覆盖面广等优点，而成为人们获取信息资源最重要的来源。

利用 TCP/IP 协议和超链接技术而形成的客户机/服务器模式的 WWW 信息系统，是不同领域、不同学科、不同性质和不同种类的信息资源的集大成，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信息资源，堪称世界上最大的图书馆或信息资源库。WWW 信息服务的特点是在网上进行多媒体信息的收集、分类、存放、发布和交流，并向网上的用户通过搜索引擎、网络指南、学科导航等方式提供信息检索服务，同时通过电子邮件、电子公告版、聊天、博客等方式向用户提供交互式服务。

1.5 信息资源与知识创新

创新是信息社会的重要特征，是当代科技、经济、社会发展的关键。而知识创新作为国家创新体系的重要组成，是知识经济时代所有创新活动的共同实质，这种创新的动力在于未来的竞争要求，创造新思想并转化为市场前景广阔、具有发展潜力的商品，为未来持续增长奠定基础。从狭义的角度理解知识创新是指通过学习研究获得和创造新知识的过程，它发生在知识生产、传播应用的全过程；而从广义的角度看知识创新则是指为了经济和社会利益创造、传播和利用新知识，使其转变成市场上的产品或服务，包括科研获得的新知识、新知识的传播和应用、新知识的商业化等。从上述界定可知，知识创新贯穿于信息活动的整个过程，因此任何知识创新活动都离不开信息的有效获取与积累、挖掘与利用。

1.5.1 信息的有效获取是知识创新的基础

从本质上看，知识创新是知识要素间的重组与发展，这种知识要素部分源自于知识创新者的既有积累，而部分则需从外部信息源获取。如果不能有效获取创新所需的信息与知识，知识创新就不可能实现。而在信息的检索过程中，人们经常感到由于信息的存储过于庞大和无序，而难于有效的获取知识创新所需的信息。因此人们需要认真学习和研究获取