

Academic Information Sources on the Internet  
—Retrieval & Analysis

# 网络学术信息资源

## 检索与利用

华薇娜 编著

国防工业出版社

# 网络学术信息资源 检索与利用

华薇娜 编著

国防工业出版社

·北京·

## 内 容 简 介

网络的发展大大扩充了我们所拥有的信息资源,“秀才不出门,能知天下事”已成为现实。然而网上的信息浩如烟海,许多人常常感觉到通过网络收集信息,绝大部分时间都花在了找信息上,但大部分时候却得不到想要的结果。网上究竟有哪些信息资源,它们在什么地方,如何在网络信息世界中遨游而又不迷失方向,这些已成为各方关注的话题。对学术界、工商企业界、政府机关、技术部门以及广大学生用户来说,最重要的是要发现、了解,并不失时机地利用那些有利于研究、开发、决策、教学等方面的网络学术信息资源。本书根据最新的网络信息资源的实况,系统地介绍综合性的和各学科网络信息资源的种类、来源、沿革、特点、使用方法,并推荐了一些重要的站点。此外,还有对网络信息分析利用方法的具体介绍。本书对于培养广大网络用户检索和利用网络学术信息的能力,提高用户在网络环境下的生存与学习能力,具有较高的实用价值。

本书可作为高等院校网络信息资源检索课程的教材,同时也面向各专业科研人员、高等学校的教研人员、IT业研发人员、市场开发人员、商务决策人员、学术或专业图书馆员、信息服务从业人员,以及广大因特网用户。

### 图书在版编目(CIP)数据

网络学术信息资源检索与利用/华薇娜编著. —北京:国防工业出版社,2002.12  
ISBN 7-118-03078-3

I. 网... II. 华... III. 计算机网络—情报检索  
IV. G252.7

中国版本图书馆 CIP 数据核字(2003)第 008609 号

国防工业出版社出版发行

(北京市海淀区紫竹院南路 23 号)

(邮政编码 100044)

新艺印刷厂印刷

新华书店经售

开本 787×1092 1/16 印张 17 1/2 406 千字

2002 年 12 月第 1 版 2002 年 12 月北京第 1 次印刷

印数:1—4500 册 定价:22.00 元

(本书如有印装错误,我社负责调换)

# 前 言

1984年教育部作出的在全国高校普遍开设“文献检索与利用”课程的决策具有战略意义，为大学生们在现代社会取得成功打下了坚实的基础。文献检索课从最初完全传统的印刷型检索工具开始，20世纪80年代后期逐步结合计算机联机检索的知识，90年代中后期又陆续融入脱机的光盘检索的内容，到如今网络信息资源的检索，在短短的十多年中发生了巨大的变化，这也许是国内教学内容变化最快的课程之一。然而，当我们果真面临汹涌澎湃的网络信息大潮时，却发现原有的教材已跟不上网络发展的步伐。我作为最初参加文献检索课程培训的第一批教师中的一员，亲身经历了这门课程的几乎各个发展阶段。现在，无论是从教材和教法的角度，这门课程所面临的形势都相当严峻。我们应该主动迎接挑战，以使本课程适应时代发展的要求。这便是本书产生的客观需要。

2000~2001学年，我有机会作为访问学者赴美进修。以一位检索课教师应有的敏感，我试用了大量的网络数据库，这大大地开阔了我的眼界，也使我感觉到有责任将国外更多的、更专业的数据库介绍给国内的同行，介绍给受教育的对象。这又是本书产生的主观冲动。

快节奏的生活步伐已使我们很难将充满活力的学生留在检索室里查找翻阅工具书，尤其是在手检工具书与机检数据库并存的情况下。因此，就信息检索类的课程而言，是否可以考虑，对于非信息管理专业的学生只讲授网络学术信息资源。尽管这还是学术界讨论的问题，但我们的学生不能等，因特网的发展更容不得我们等，编写适合于网络学术信息资源教学的教材再一次被客观需要提上了日程。况且网络信息资源发展之迅速已超出人们的想象和预测，二年级学生在检索课上所学的内容到学生毕业前做论文时有的已发生了很大的变化。因此，需要有适合于广大网络信息用户的自学用书做补充。于是，本书就在这样的背景下、这样的原则下应运而生了。

本书基于我近年来在网络信息资源方面的学习与研究，书中的基本内容经过了南京大学本科生的全校公选课、辅修课和研究生课程的教学实践。所有的网络资源在成书前的几个月内经过多次检验调整，以保证内容的新颖性。本书网络资源的范围从国内到国外，侧重于国外的数据库与网站，但同样重视国内的综合性、专业性资源。网络资源的可利用程度从收费网站到免费网站，在数量上以免费网站为多。免费网站的收集一般以国外研究单位推荐介绍以及专业评价单位推选为最佳网站的结果为据；同时也介绍重要的以及正在发展中的收费网站，以便人们了解、选择或引进。在本书中凡未注明收费资源的，一般均为免费资源。网络资源的学科范围从人文科学、社会科学到自然科学，涉及学术研究的诸多领域；既有全文数据库，也有目录性数据库，在数量上以后者为多；篇幅上有详有略，重要的数据库除了介绍其特征及用途外，还选择性地介绍了其发展历史、检索方法等，一般

的数据库都有具体收录内容的介绍。此外，每个数据库都特辟专门章节介绍网络信息资源的检索技巧及分析处理结果信息的一般方法。

因特网确实是一个获取学术信息的极有价值的来源，但同时又确实存在着大量的无用信息或信息冗余，因此我们首先必须了解它、熟悉它。本书从高层次信息需求的观点出发，系统地介绍了学术性、研究性信息的网上检索和获取的方法、途径，以及一些相关的重要站点，以帮助人们准确地选择适当的数据库，迅速地发现并了解所需要的信息资源，全面地获取第一手信息。也许这些资源分散在世界各地，或躲藏在天涯海角，但对一个掌握了网络学术信息检索真谛的用户而言，只需在你的办公室或实验室或宿舍里轻轻点击鼠标，你就可以得心应手地将它们收集在一起，归拢到你的文件夹中来。这个时候，你一定会感受到信息全球化的美好。

本书在成书过程中，得到信息检索方面的老前辈陆宝树先生的指导，也得到南京大学信息管理系师生诸多方面的支持，还得到南京大学许多专业院系学者们的帮助，在此一并表示感谢。总装备部档案馆郭建华同志参加了本书的部分编写工作。由于初次尝试，加之工作量大，时间紧，错误之处敬请读者批评指正。

曾有一位准备参加研究生复试的学生希望能比较全面地找到他所投考的某位导师的论文，在图书馆折腾了多天，仍没能满足自己的要求，于是向我求助。我告诉他，只需几分钟便能解决问题，甚至都可以不用走出他的宿舍，所需结果就在手指的几次点击之下。这位学生后来告诉我，自己几乎每天上网，却不知各种数据库的用法，似乎有种出入豪华餐厅，却只会吃水煮青菜的感觉。如果说网络信息资源就像一席豪华的自助餐的话，你若不去找菜单，不知有些什么，哪儿有，怎么吃，那恐怕就只能吃点水煮青菜了。对于广大的网络用户而言，网上不仅有聊天室，有游戏玩，有 MP3 可欣赏，也不仅只有搜索引擎作为检索的途径；网上有着更为丰富精彩的、有助于学习、研究、开发的学术信息资源宝库。本书就是在这个宝库的一角为你开启的一扇小窗，从这里探个头进去吧，看看里面的世界有多精彩。

华薇娜

2002年10月于南京大学大钟新村宿舍

# 目 录

<b>第一章 网络信息资源概述</b> .....	1
第一节 网络信息资源的概念.....	1
第二节 网络信息资源的种类.....	1
第三节 网络信息资源的特点.....	4
第四节 网络信息资源的共享.....	5
第五节 网络信息资源的划分与评价.....	7
第六节 网络信息资源检索的一般方法与技巧.....	9
<b>第二章 搜索引擎</b> .....	13
第一节 搜索引擎概述.....	13
第二节 搜索引擎实例.....	14
<b>第三章 综合类网络信息资源</b> .....	28
第一节 国内主要数据库系统.....	28
第二节 国外主要数据库系统.....	50
<b>第四章 工具书类网络信息资源</b> .....	68
第一节 综合性工具书类网络信息资源.....	68
第二节 百科全书类网络信息资源.....	72
第三节 词典类网络信息资源.....	75
第四节 人物和机构名录类网络信息资源.....	80
第五节 书目类网络信息资源.....	84
第六节 地名地图类网络信息资源.....	88
<b>第五章 艺术、人文科学类网络信息资源</b> .....	92
第一节 人文科学综合性、跨学科类网络信息资源.....	92
第二节 艺术类网络信息资源.....	96
第三节 语言文学类网络信息资源.....	105
第四节 历史类网络信息资源.....	114
第五节 哲学类网络信息资源.....	118
第六节 宗教类网络信息资源.....	120
<b>第六章 社会科学类网络信息资源</b> .....	123
第一节 社会科学综合性、跨学科类网络信息资源.....	123
第二节 政治学网络信息资源.....	126
第三节 经济、商务类网络信息资源.....	130
第四节 法学、法律网络信息资源.....	147

第五节	社会学网络信息资源	154
第六节	新闻、传播、大众传媒类网络信息资源	157
第七节	心理学网络信息资源	161
第八节	教育类网络信息资源	164
第九节	图书馆学情报学网络信息资源	169
<b>第七章</b>	<b>自然科学类网络信息资源</b>	<b>174</b>
第一节	科学技术类综合性跨学科网络信息资源	174
第二节	数学类网络信息资源	177
第三节	物理、电工电子、计算机类网络信息资源	184
第四节	化学类网络信息资源	192
第五节	地学与环境科学类网络信息资源	199
第六节	生物学网络信息资源	205
第七节	医学类网络信息资源	210
第八节	农业类网络信息资源	220
<b>第八章</b>	<b>工程类网络信息资源</b>	<b>225</b>
<b>第九章</b>	<b>特种文献类型网络信息资源</b>	<b>238</b>
第一节	会议类网络信息资源	238
第二节	学位论文类网络信息资源	242
第三节	科技报告类网络信息资源	245
第四节	专利类网络信息资源	249
<b>第十章</b>	<b>网络学术信息资源的收集技巧与分析利用</b>	<b>256</b>
第一节	网络学术信息资源的收集技巧	256
第二节	网络学术信息资源的分析利用	258
<b>附录一</b>	<b>南京大学学生信息分析成果题录</b>	<b>266</b>
<b>附录二</b>	<b>留学类网络信息资源</b>	<b>268</b>

# 第一章 网络信息资源概述

## 第一节 网络信息资源的概念

信息资源、物质资源和能量资源构成现代社会经济发展的三大支柱，有效地获取信息资源已成为推动社会发展和企业进步的重要力量。因特网的出现与发展是人类文明史上的重大事件，它对科技发展、经济繁荣和人类进步起着重要的推动作用。

网络信息资源是指以电子资源数据的形式将文字、图像、声音、动画等多种形式的信息存放在光、磁等非印刷质的介质中，并通过网络通信、计算机或终端等方式再现出来的信息资源的总和。

因特网的发展，让人们感到，一方面，相对于传统的信息获取方式，我们能以比以前快得多地速度在全球范围找到我们所需要的信息；另一方面，因特网的发展大大扩展了我们所拥有的信息资源世界，信息的生产和服务都能够直接面向用户，因特网成了世界上最大的信息集散地和最大的图书馆。然而，人们很快发现，因特网就如同一个将所有的物件都堆在地板上的豪华别墅一样，既丰富多彩又杂乱无章。要想在这个变化多端又鱼龙混杂的信息海洋中发现并找出具有利用价值的信息并不是一件容易的事。以致于许多人感叹，通过因特网收集信息，绝大部分时间似乎都是在找信息，而不是找到信息。网络用户如果不了解网络信息资源的分布与规律，再丰富的网络世界又有什么意义？对于网络用户而言，因特网上究竟有哪些可以被利用的信息资源，它们是如何分布的，有什么特点，如何正确选择合适的那一部分网络信息资源库并制定正确的方案从中迅速找到所需要的信息，而不至于在其中淹没和迷失，这些都是在进行网络信息资源检索之前值得考虑的一些问题。

## 第二节 网络信息资源的种类

网络信息资源包罗万象，广泛分布于整个网络之中，既没有统一的组织管理机构，也没有统一的目录。若果真要划分的话，也有多种划分标准，现从中总结出如下一些结果：

### 一、按照所采用的网络传输协议划分

#### 1. WWW 信息资源

WWW(World Wide Web, 中文名称为万维网, 也称环球信息网)是 20 世纪 90 年代初期由位于瑞士的欧洲研究中心发明的。由于它能方便迅速地浏览和传递分布于网络各处的文字、图像、声音、和多媒体超文本信息，并适用于因特网信息服务，因此在 90 年代



中后期得到迅速发展。因特网上的 WWW 服务器以每年翻几番的速度增长，成为因特网信息资源的主流。

### 2. FTP 信息资源

FTP (File Transfer Protocol) 称为文件传送协议，是因特网上历史最为悠久的网络工具。它允许人们通过协议连接到因特网的一个远程主机上读取所需文件并下载，换句话说，它相当于在网络上的两个主机间拷贝文档。因特网刚开始流行时，网上文件大部分都是 FTP 站点上的，因而 FTP 在因特网的发展史中发挥着重要的作用。至今 FTP 仍是发布、传递软件和长文件的主要方法，而且许多文件在 FTP 服务器上，并没有做 http 的链接，值得人们重视。

### 3. Telnet 信息资源

Telnet 是因特网的远程登陆协议，允许用户将自己的计算机作为某一个因特网主机的远程终端与该主机相连，从而使用该主机的硬件、软件和信息资源。

### 4. 用户服务组信息资源

用户通信或服务组是因特网上颇受欢迎的信息交流形式。其中包括新闻组 (Usenet Newsgroup)、电子邮件群 (Listserv)、邮件列表 (Mailing List)、专题讨论组 (Discussion Group) 等。它们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，是因特网上进行交流和讨论的主要工具。它们的工作原理与使用方法也非常相似，均用于网络用户间的信息交流；但又各具特色和用途，锁定各自特定的用户。USENET 是因特网上的一种应用软件，用于提供新闻组服务。在这个服务体系中，有许多新闻服务器，用户可在自己的主机上运行新闻组阅读器软件，申请加入某个新闻组，并从服务器中读取新闻组消息或将自己的意见发送到新闻组中。用户可查阅别人的意见并予以回复，由此反复，进行讨论，形成一种特别的信息资源。电子邮件群和用户邮件组有管理地址和邮件地址。通常电子邮件群在网络地址前加“LISTSERV@”，用户邮件组则在网络地址前加“\_request@”。加入电子邮件群和用户邮件组必须申请订阅，待得到同意的回复后才能通过其发送和接收信息。

### 5. Gopher 信息资源

Gopher 又称信息鼠，这是一种基于菜单的网络服务，类似万维网的分布式客户机/服务器形式的信息资源体系。它是因特网上一种分布式信息查询工具，各个 Gopher 服务器之间彼此连接，全部操作都在一级级菜单的指引下，用户只需在菜单中选择和浏览相关内容，就完成了对因特网上远程联机信息系统的访问。此外，Gopher 还可提供与前文所提及的其他多种信息系统的连接，如 WWW、FTP、Telnet 等。

## 二、按照信息资源的有偿性划分

1. 收费类信息资源：用户需要先行注册登记，并通过一定的付费方式交纳所需费用后方可使用。但对于局域网内用户来说，一旦主机系统成为某系统的注册用户后，所有网内用户即可登陆该系统免费使用其数据库，不过有些系统会控制局域网内并行用户的数量。一般来说，商业性机构开发的大型的、历史悠久的、成本费用高的系统采用收费方式的较多。当然也有一些很有赢利性的系统出于公益等目的，放弃收费。

2. 免费类信息资源：这是一些无偿服务的数据库，多为非赢利性公司开发，或小型

的、实验性的、新上网的数据库，多数不需要注册，用户可直接连接使用，但也有些系统虽然免费，却要求用户注册，以了解、掌握用户信息。

然而这两类数据库不是固定不变的。为了让更多的用户了解并成为其数据库的购买者，收费数据库常通过组织免费试用的方法，向各科研教学单位推广介绍其数据库。因而试用阶段的收费数据库一般是免费的。反之，相当数量的免费数据库在得到一定数量的用户后，又转变为收费数据库了。但一般来说，许多收费的数据库同时又为新用户开放一定时间范围的试用期（Free Trial），用户通过注册后，可在试用期内免费试用数据库的全部或部分内容，国外的一些大型的或重要的数据库尤其如此。

### 三、按照信息资源的内容划分

1. 学术研究类信息资源：针对大学、研究所以及图书馆所设计的数据库，大多是在有着悠久历史、有着大量用户的印刷版的工具书或检索刊物的基础上开发，以检索目的为主，且多为收费数据库。

2. 教育类信息资源：针对各类学校，尤其是中、小学的学生、教师以及家长开发的数据库，以教育目的为主。很多数据库联系学校开设的各类课程，并结合学生用户的特点设计页面，且多为免费数据库。

3. 政府信息资源：由各级政府机构、部门开发，向公众公布政府信息或用于政府部门间信息的交流，是电子政府的重要组成部分。这类信息资源既有全公开的，也有只对特定对象公开的；有收费的，也有免费的。

4. 商业经济类信息资源：这是一种为特定行业、特定对象服务的网络信息资源，如各类机构名录、产品目录、经济贸易、市场、金融证券等，有收费的，也有免费的。

5. 生活娱乐类信息资源：这是为普通公众服务的信息资源，也是内容最庞杂、形式最多的一类网络信息资源。

6. 广告信息资源：出于经营的目的而设计，一般不单独存在，而是附属于其他的网站，尤其是大量的免费网站。

### 四、按照信息资源的加工程度划分

#### 1. 一次资源信息

是指那些未经加工的、不具备检索入口的原始资源信息。在网络信息资源中多为个人网站上的信息。这些信息从内容上来看一般比较详细、准确，但分散无序。

#### 2. 二次资源信息

将分散的、无组织的依次信息，根据一定的方法，经过整理、归纳、简化，把可供检索信息的特征著录下来而形成的有组织、有系统的信息类型。这类信息提供获知一次信息的渠道，并提高一次信息的传递速度和使用效率，多为文摘索引类的具有标引检索性质的网络信息。

#### 3. 三次资源信息

在一、二次信息的基础上，通过评价筛选，以简练的文字编写而成，提供检索入口并能被用户直接利用的信息资源，多为工具书类的网络信息。

#### 4. 全文信息

若根据信息的责任者来判断,这也是一种原始资源信息。但从加工程度上来看,这却是经加工了的、在数据库中供全文检索的信息。全文检索是指以全文本信息作为检索对象,建立全文数据库,除了具有布尔逻辑检索功能外,还具有文本检索功能,并允许用户以自然语言检索,直接获得原文或原文中相关的章节、段名等。全文检索具有全新、方便的检索功能,可以直接根据文献资料的内容进行检索,支持多角度、多侧面的综合查询方式。

### 第三节 网络信息资源的特点

了解网络信息资源的特点,是为了清楚地知道网络信息资源的优势与劣势,在网络信息资源的检索与利用方面,更好地发挥主观能动性,做到正确地扬弃。

#### 一、信息量大,传播广泛

网络信息资源极为丰富,因特网已经成为继电视、广播和报纸之后的第四媒体。它既是信息资源存储和传播的主要媒介之一,也是集各种信息资源为一体的信息资源网。据估计,因特网每天发布的信息超过 700MB 以上。因特网每天发布的信息总量约为 450MB,全网信息总量达 20TB。由于信息源的增多,信息发布的自由,网络信息量呈爆炸性增长。几乎只要有电脑的地方,就有信息的传播。

#### 二、信息层次多,品种多样

因特网上的信息资源层次众多,有一次信息、二次信息、三次信息;有文本信息、图像信息、图形信息、表格信息、超文本信息等,包括各种电子书刊、书目数据库、联机数据库、软件资源等,是多媒体、多语种、多类型信息的混合体。网络信息资源包罗万象,几乎覆盖各学科、各领域、各地域、各语言的信息资源;信息发布者既有政府部门、高等院校、科研院所、学术团体、行业协会,更有大量的公司企业和个人。

#### 三、自由发布,交流直接

除了以往联机检索以及大量在图书馆工具书、检索刊物的基础上发展起来的数据库这些正式交流渠道的信息交流渠道外,网络信息资源中更多的是非正式交流渠道发布的信息。这里提供了自由发表个人见解的广阔空间和获取非出版信息的丰富机会,包括那些正式出版物中所不能得到的信息,如类似于灰色文献的信息、还未成熟的观点、个人网页上的研究心得、教学资料等。同时,因特网扩大了人际交流的空间,如新闻组、讨论组、邮件列表等,都为用户提供了更多的直接交流的机会。

#### 四、信息传播速度快,变化频繁

在非网络信息中,信息传播速度快且变化最大的莫过于报纸,但是报纸一经出版,其信息便无法更改,而在因特网上,信息的更新相当及时,不少新闻站点、商业站点的信息每日更新,信息传播速度十分快,信息能瞬时实现交流。

## 五、检索方便, 价廉实惠

在友好的用户查询界面下, 用户可以根据需要和已知信息任意选择检索方式与入口, 进行自然语言检索、全文检索, 方法简便易学。由于输出的信息按相关性大小顺序排列, 通过关键字词的查询, 可迅速找到所需内容; 此外用户还可以在更大的库存容量范围、更大的时间跨度范围检索, 且有可能在更多的数据库集合中总体检索, 因此检出的信息更对口、更专指、更全面。这些都是网上信息资源相对于传统信息资源的最大优势。又因其不需批发、代理、零售, 没有中间环节, 大大降低了成本。

## 六、分散无序, 缺乏管理

网络信息资源的分散表现在信息没有一个中心点, 也没有全面性的权限, 甚至连接本身的意义也显得模糊和多样, 通过一种文献可以连接到更多相关或相似的文献; 同样, 这份文献也可能是从另一份文献连接而来的, 这种前所未有的自由度使网络信息资源的共建和共享变得潜力无穷, 然而也使信息资源处于无序状态, 而且, “海量”的信息和快捷地传播加剧了网络信息的无序状态, 许多信息资源缺乏加工和组织, 只是时间序列的信息堆积, 缺乏系统性和组织性。网络信息的地址、连接、内容等都处于经常性的动态变化之中, 其变化、更迭、新生、消亡等都时有发生, 没有统一的管理机构来控制; 信息的发布有很大的自由度和随意性, 也没有统一的标准, 缺乏必要的过滤、质量控制和管理机制。

## 七、内容庞杂, 质量不一

因特网的信息大都没有经过严格的审查, 信息发布具有很大的随意性和自由度, 信息缺乏必要的过滤监督和质量控制。信息内容十分庞杂, 正式出版物与非正式出版物交织在一起, 科技信息、学术信息、商业信息、个人信息与一些暴力、色情等不健康信息混为一体, 既有大量国际水平的研究成果, 又有许多难登大雅之堂的涂鸦之作和虚假信息, 信息质量良莠不分、参差不齐, 给利用有价值的网络信息带来极大的不便。

# 第四节 网络信息资源的共享

网络信息资源作为一种新兴资源, 在信息资源中占有越来越大的比重, 对社会的发展和进步起着不可忽视的作用。网络信息资源无限共享模式是以信息资源在网络中的跨时空广泛共享为出发点进行网络信息资源建设的, 从而实现网络信息资源在最广大用户范围内充分开发利用。网络信息资源无限共享不仅是一种信息资源生产者、加工者、服务者与用户之间的广泛共享, 而且是一种真正的信息共享, 不再是文献载体的共享。

## 一、网络信息资源的共享的可能性

实现稀缺资源的共享是人们长期以来的共同理想。信息资源虽然数量庞大、类型多样, 且不易损耗, 但它仍然是有限的。在传统的技术条件下的信息资源共享主要是信息

物质载体的共享,即文献的共享。这种共享由于受共享交易成本的限制,只能是在一定时间、空间范围内有限的共享。而计算机网络的兴起为信息资源共享带来了新的契机,开辟了信息资源无限共享的新天地,是网络信息资源共享的重要条件。

计算机网络从本质上说是一个宽带、高速、动态的电子信息网络,它能够随时随地给用户大量的信息资源。数字化技术使信息的复制变得更加简单,网络传输信息能瞬间到达遥远的地方,无线通信能轻松地跨越地域的障碍。这种相互连接的网络是传统时空概念的破坏者,它将时间与距离几乎缩小为零,使世界成为一个地球村。正是由于网络这种无限扩展的特性,使每个国家、单位、个人都有可能联入网络,实现跨时空的信息资源共享。所以网络信息资源共享也是缩小我国与发达国家差距的捷径,是我国实施科教兴国的重要手段。

图书馆网络化,实现文献资源共建共享已成为一个国家图书馆事业发展水平的重要标志,也是社会进步、经济繁荣的具体体现。同时网络化是实现图书馆现代化的重要环节。图书馆网络化可以实现文献资源的远程化传递,大大提高馆藏文献的利用率和馆藏文献的拥有量,真正实现整个人类社会文献资源的共享,满足社会各界的文献信息需求,更好地履行图书馆的社会职能。在网络环境下实现图书馆的资源共建共享的社会需求是网络信息资源共享成为可能的重要因素。

## 二、网络信息资源共享中存在的障碍与问题

### 1. 信息庞杂

随着当代科学技术的迅猛发展和各学科间的渗透、交融,科学出版物的数量也急剧增加,信息“爆炸”处于失控状态。虽然信息技术的迅猛发展,使文献信息上网时间缩短、传播速度加快,数字化技术也使信息的复制变得更加简单,但是由于人类创造的知识成果分布在不同语种、不同历史时期、不同出版物类型和不同载体中,再加上文献信息资源入网尚不规范,网上信息十分庞杂,加大了知识信息的获取的难度。

### 2. 信息污染

网络是一个开放的系统,每个人都可以在网上发布信息,而任何个人和团体都难以干涉他人的网上行为。网上的这种自由性和开放性,造成了信息传递的无序和失控,使得互联网上的信息海量剧增,不可避免地出现了垃圾信息、虚假信息、冗余信息等,影响了信息的传输速度,侵占了大量的网络资源和用户查询、检索信息的时间,从而给用户带来了财力、物力和精力的浪费。同时由于大量污垢信息存在,使用户对真实信息的信任程度大为降低,影响了有效信息的利用。

### 3. 缺乏宏观调控

我国的网络信息资源虽然发展迅速,但仍缺乏相应的管理体制,宏观协调控制的能力不强,许多信息资源建设都是各自为政。由于缺乏分工合作,使网络信息资源大量重复建设,造成人力、物力、财力的巨大浪费。同时,由于资源的分散,不能集中力量进行重点建设,使我国的网络信息资源建设质量水平不高,产业化程度低下。

### 4. 狭隘的资源共享观

在传统技术条件下的资源共享是文献载体的共享,而文献的异地传播需要花费相当多的时间和费用,当共享的这种交易成本超过预期收益或文献定价的时候,共享行为就

不会发生,所以此时的信息资源只是在一定区域和部门中共建共享。传统技术条件下的资源共享观念在网络环境下已不再适用。在网络环境下,信息异地传播的成本将极大降低,而传统的资源共享观念远没能发挥出网络跨时空资源共享的优势。

#### 5. 规范化、标准化程度不高

没有规范的接口和统一的技术标准,资源共享将成为空谈。现在基于因特网的 WWW 信息资源基本已实现了标准化,共同遵守 TCP/IP 协议。但其他网络资源,尤其是数据库的标准化进程却很迟缓。数据库的应用受到严重的限制,许多数据库不能相互交流或交流的范围有限。我国在文献分类、编目标标准化方面曾经作过很多努力,也取得了很大的进展。但是在网络信息资源标准化建设方面已经开始落后,理论研究开展的很少,实践中执行标准化建设方案的更少。很多图书馆还是各自为政,信息资源建设和服务对象自成体系,这与网络信息资源开放的特性格格不入。

#### 6. 网络通路不畅

网络通路不畅是指计算机信息网络在运行过程中由于路窄、人多、入口少而出现的通路不畅或不能进入的现象,形成通路障碍。目前有些网络的节点拥有较多的拨号上网户,而可利用的电话号码数比上网户数少得多,一旦所有的电话号码被占用,其他上网户就被排斥在网络之外,这就使网络用户不能及时上网获取所需的网上信息资料。通路障碍轻则降低网上信息的传递速度、减少单位时间的信息获取量而提高信息成本;重则造成通路堵塞,不能在一定时域内获取网上信息,降低信息的使用价值,客观上对网上信息的共享形成障碍。

### 三、注意网络信息资源的合法利用

网络空间绝不是非法使用版权作品的天堂,现行著作权法中的某些条款、知识产权界的理论探讨,尤其是国内外司法实践已反映出网络环境下版权保护的趋势与模式,为合法使用网络信息资源、防止侵权提供了依据。未经许可擅自传播版权作品(包括下载、上载、网间转载等)、未经授权的连接(包括连入、镶边等)、破坏版权管理信息、破坏技术保护措施等都属于网络信息资源的侵权形式。此外,还要注意那些不易察觉的侵权行为,要了解到如没有版权标记的,不一定不受版权保护;使用微量的版权作品,也有侵权危险;付款不能代替许可;网上作品是受著作权法保护的。总之,人们在充分开发网络信息资源的同时,要注意遵纪守法,也只有如此,才能真正充分地利用网络信息资源。

## 第五节 网络信息资源的划分与评价

目前,因特网上的信息量已经以海量状态存在,人们面临的最大问题已经由信息短缺转化为信息过滥。因此,要使得网络信息检索更为高效有序,使检索者能在尽可能短的时间内查找到所需信息,对网络信息资源的正确划分与评价就变得愈来愈重要。

国内外都有许多对网站划分评价的实例,但至今还没有形成统一的标准。这里不妨举少量实例供人们参考。

### 1. Dow Jones Business Directory <<http://bd.dowjones.com>> 的评价标准

它为商业网站设计的评价标准分为四大项,即内容、速度、导航和设计。每项得分从1分到10分不等,最高得分可以是40分,但获得这个分数的网站极为少见。

### 2. Britannica Internet Guide <<http://www.eblast.com>> 的评价标准

这是一个著名的百科知识网站。它列出了五种考评标准,即信息的精确度、深度和宽度,出版者的权威性、版面设计、图表、多媒体效果,导航系统的简明易用性,内容更新速度。除此之外,所收录的网站还要接受从“有收录价值”到“最优”的五个级别的评定。

3. 由 ArgusAssociates 公司制作的信息评价工具 ArgusClearinghouse, 根据站点的资源描述、主观评估、设计水平、组织机构和资源指南本身等指标进行评价。

4. 澳大利亚国立大学从1994年起,每年进行一次,从站点的内容质量、组织机构、外观设计等方面,对一些学术网站进行评价。

5. 比较著名的 Lycos Top 5% of the Web 将网络信息分为新闻、旅行、科学、文化、商业、时尚、体育、技术、教育、购物、娱乐、政府、金融、保健、生活方式、儿童、职业、汽车等18个类型,经过专业评价人员的分级评价,收集每个类型评分较高的网站。评价指标为内容、设计和总体表现。

6. 美国依阿华州立大学的 Cyberstarcks 站点,主要评价科学、技术方面的网站,评价指标为内容的权威性、准确性、清晰程度、独特性、新颖性、相关评论和社区需要等。

7. SelectSurf 的评价指标是内容、适用性、设计和易用性。对评价后的网站给出一段简单的评语,并标注黄色记号。

### 8. 美国图书馆协会的评价标准

美国图书馆协会(American Library Association, 简称: ALA)下属的参考馆员与用户服务协会(Reference and User Services Association, 简称: RUSA)在1998年ALA年会上其下属的机助参考部(Machine-Assisted Reference Section, 简称: MARS)任命了一个特别小组,建立一个方法与任务系统评定万维网上突出的参考网站。至今为止,该组织已连续三年评出了万维网上的最佳免费网站。MARS 所采用的最佳网站评价标准包括如下几方面:①质量、深度和内容的有用性;②便捷参考;③方便使用;④内容的即时性;⑤好的用户服务系统;⑥制作者的权威性;⑦内容的独特性;⑧有效性;⑨有效地运用新的技术手段。

### 9. 国内学者的总结

除了国外网站划分的一些标准外,国内亦有学者提出各种评价标准。如郑睿提出下述七个方面的评价指标,即:功能、设计、内容、原创性、专业程度、查检性、整体印象等。陆宝益则提出了包括18个项目或方面的评价标准,即:①信息内容的深度与广度;②引用数据或事实的准确性;③表达观点的客观性;④创新性;⑤稳定性;⑥安全性;⑦优秀的导航系统;⑧责任者的身份和学术地位,包括团体责任者的学术地位、经费、组织方式、运行机制等;⑨信息来源或提供商的权威性;⑩注释或参考文献是否规范、统一、完整、准确;⑪创作(办)目的;⑫用户界面友好性;⑬外观设计;⑭对用户的技术要求;⑮媒体形式;⑯更新频率;⑰专业信息比例;⑱信息组织层次。

尽管各人评价标准不一,但有一点可以总结,即好的参考网站应是那些检索方法易于检索者理解,能够供检索者在特定范围内全面、迅速、准确地查得结果,并以用户所需格式提供答案的站点

对于评价和选择网络学术信息资源,还有下列几方面因素需要强调,即:

#### 1. 重视内容准确性 强调学术价值

高校师生和研究人员查找利用信息主要是为满足教学科研的需要,对信息质量要求较高。应选取某学科范围内有学术价值、有一定深度、能反映本学科的前沿发展水平和动态的网上学术资源。学科的内容范围和准确性应是首先考虑的重要指标。

#### 2. 重视信息制作发布者的可信度

选择印刷版图书时,著者、出版社是一个重要参考因素;选择期刊时首选核心期刊;而网上信息的制作发布者是网络信息资源的一个重要考虑因素。权威信息中心或情报机构、本学科学术刊物的出版单位、各种社会组织(政府、大学、研究机构、图书馆)所制作发布并提供上网的信息是学术性信息的主要来源。

#### 3. 重视信息的稳定性

网络信息资源呈动态变化性,而网站、网页形式相对稳定有利于用户使用。印刷型文献的数字化、网络期刊、联机数据库、图书馆目录等都是比较稳定、准确可靠、方便存取的信息资源。

#### 4. 使用的方便程度

教学科研任务的前沿性要求使高校师生必须面对数量极大的最新专业文献,因此网站能否方便使用,是否符合专业人员查找相关文献的习惯,是否允许多种访问工具,是否能在较短的时间内进入并搜索到所需的最新资料是应考虑的因素。

## 第六节 网络信息资源检索的一般方法与技巧

信息在网络上的急速增长,引起网络信息的生产和利用之间的尖锐矛盾,一方面是网上存在大量的网络信息,而另一方面是人们利用网上信息资源愈来愈困难。对一般用户而言,他们并不了解网络检索方法和技巧,以致检索的结果要么与所要内容不相关联,要么检索不到所需的内容。如何在网上全面、准确、快速并且没有重大遗漏地获取有用的信息便成为现代人们,尤其是学术科研人员所面临的一个重要问题。只有了解网络信息检索的一般方法与技巧,充分利用网络提供的检索条件,制定合适的检索策略,才能在网络这个茫茫大海中,巧妙地利用各种导航,控制手中的鼠标如同驾驭着检索的航船,驶向既定的目标彼岸,并最终发现新大陆。

基于现代信息技术的网络信息检索突破了时间、空间的限制,为我们提供了传统的手工检索阶段不可能实现的,甚至不可想象的检索功能,但其检索原理却在很大程度上传承了手工检索的基础,这有助于人们在全面地继承原有检索理论的基础上不断地扩展和创新。要想取得满意的检索结果,以下一些基本的检索命令或符号是在构造检索逻辑式时所需要掌握的。



## 一、布尔逻辑符号的使用

布尔逻辑检索是检索系统中应用最为广泛的检索技术,是最简单、最基本的匹配模式,几乎所有的网络信息检索系统都支持布尔逻辑检索。

基本布尔逻辑运算式有三种:

逻辑与(and),在布尔表达式中通常记为“\*”,也有的系统中用“&”表示。逻辑与可以检索出同时含有检索词A检索词B的信息,用于交叉概念或限定关系的一种组配,可缩小检索范围,提高查准率。如 Color AND Television,表示要求检索“Color Television”的结果。

逻辑或(or),在布尔表达式中通常记为“+”,有的系统中用“|”表示。逻辑或可以检索出含有检索词A或者含有检索词B或者含有A、B的信息,用于并列概念的一种组配,可扩大检索范围,提高查全率。如 Sensor OR Detector,表示要求检索“Sensor”或“Detector”的结果。

逻辑非(not),在布尔表达式中通常记为“-”,有的系统中用“!”表示。逻辑非能够检索出检索词A而不含有检索词B的信息,是一种用于从原来的检索范围中排除不需要的概念,或影响检索结果的概念的组配。如 Energy NOT Nuclear,表示要求检索“Energy”,但排除了“Nuclear Energy”的结果。

基本布尔逻辑运算的执行顺序:NOT,AND,OR,但可根据需要用括号重新规定顺序。在布尔逻辑检索中,另有布尔逻辑提炼符( )。括号的作用是使括在其中的操作符先起作用,或者说先被计算机解释,逻辑与(and)、逻辑或(or)、逻辑非(not)跟随其后。使用括号的目的是保证计算机确切知道执行的步骤。如果要检索两个以上的主题,必须告诉计算机首先执行哪一部分检索,在括号中被包括的词总是被计算机首先执行,如果检索顺序很复杂,那么在检索中使用括号是不可忽视的。

## 二、通配符的使用

也叫截词符,在进行信息检索时,可以在检索词的后面加一个通配符来进行截词检索,以代替任意的字母组合,这样就可以把包含某一词根的不同形式的资料同时检索到,减小了由于单词的不同形式造成的含有某一词根的检索资料的漏检,以提高查准率。通配符号大多选用星号“\*”或问号“?”,但有的数据库有其本身特定的通配符号,需要阅读该数据库的“帮助”界面得到准确的符号。如 econom\*,可分别得到 economy, economic, economics, economical, economist, economize 这样的结果。

## 三、near 符号的使用

有些网络数据库提供 near 检索操作符,即提供临近位置检索,其彼此间距可以通过使用 near 后的数值来控制,用于寻找在一定区域内同时出现的检索单词的文档。符号记为(N)或(nN),前者表示其两侧的检索词彼此相邻,位置可以颠倒,如 library (N) digital,其检索结果为 digital library,或 library digital,后者括号里的小写n,表示其两侧的检索词间允许插入 0~n 个词,且顺序可颠倒。例如: econim\* (2N) recovery,即查找所有 econim\*和 recovery 之间的间隔不超过 2 个单词的信息结果,如 economic recovery, recovery