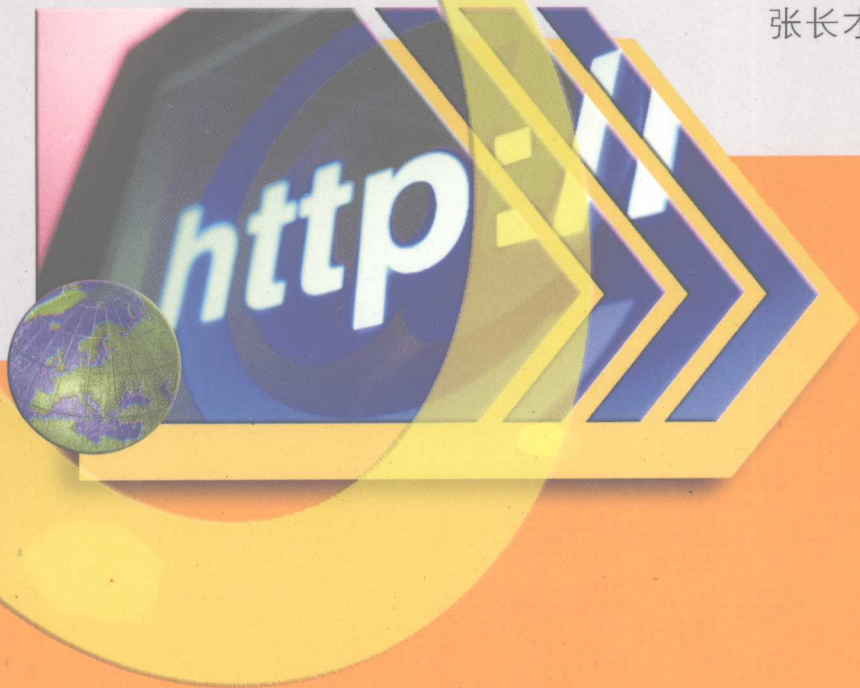


现代高等教育读本

Internet 数字信息 检索方法与技巧

主编 伦志军
李博昭
张长才



吉林人民出版社

现代高等教育读本

Internet 数字信息 检索方法与技巧

主 编 伦志军 ✓
李博昭
张长才

吉林人民出版社

图书在版编目(CIP)数据

Internet 数字信息检索方法与技巧/伦志军 李博昭 张长才编. - 长春:
吉林人民出版社,2006.9

ISBN 7-206-04749-1

I. Internet… II. ①伦… ②李… ③张… III. 文科(教育) - Inter-
net 数字信息检索方法与技巧 IV. G642

中国版本图书馆 CIP 数据核字(2005)第 082984 号

现代高等教育读本

Internet 数字信息检索方法与技巧

主 编: 伦志军 李博昭 张长才

责任编辑: 孙 一

版式设计: 刘荣泽

责任校对: 韩晓娟

出 版: 吉林人民出版社

经 销: 全国新华书店

印 刷: 长春市南关区太平彩印厂

开 本: 850 × 1168 毫米 1/32

字 数: 200 千字

印 张: 8.25

版 次: 2006 年 9 月第 1 版

印 次: 2006 年 9 月第 1 次印刷

印 数: 1-3000 册

书 号: ISBN 7-206-04749-1

定价108.00元(每册18.00元)

《Internet 数字信息检索方法与技巧》

编委会名单

主 编	伦志军	李博昭	张长才
副主编	李正红	马 戡	王 中
	武英杰	赵连堤	柳淑荣
	祖力纳	高海鹰	黄强伟
编 委	(按姓氏笔画排列)		
	田冬梅	白 岩	白淑春
	李俊杰	张月春	张永欣
	张 勇	罗 莹	周 萍
	郝志维	徐桂香	崔秀丽
	盖学良	韩 力	鲁毅巍
	霍军军		

单 位 介 绍

伦志军	吉林大学第一医院
李正红	吉林大学第一医院
田冬梅	吉林大学第一医院
徐桂香	吉林大学 图书馆
白淑春	吉林大学 图书馆
白 岩	吉林大学第一医院
韩 力	吉林省卫生统计信息中心
盖学良	吉林大学第一医院
鲁毅巍	吉林大学中日联谊医院
张永欣	吉林大学第二医院
李俊杰	东北师范大学图书馆
崔秀丽	威海新力热电有限公司
周 萍	吉林大学管理学院
张月春	吉林大学管理学院
张 勇	吉林省 图书馆
罗 莹	吉林省 图书馆
郝志维	吉林省 图书馆
霍军军	吉林省 图书馆

前 言

随着 Internet 的发展,地球越来越小,变成了一个村庄。Internet 将全球的各种信息资源带到我们每一个人面前,Internet 已成为知识经济时代不可缺少的基本工具。可以说,您所需要的绝大多数信息都可在 Internet 上获取,而且它们中多数都可以免费获取,关键在于您能否准确地找到它们。为了帮助您检索和获取丰富的网络信息,很多网络信息检索工具应运而生,如普通检索工具和专业检索工具。但是,令您苦恼的是,即使使用这些检索工具,您往往得到的是成千上万条似是而非的检索结果。您如何能在最短的时间内,选择到最合适的检索工具、使用最准确的检索方法,检索到您最需要的信息?这正是编写本书的目的。

本书遵循“授人以鱼,不如授人以渔”的宗旨,不打算简单罗列那些随时都在变化和过时的网址,不打算机械地、肤浅地介绍所谓流行网站或工具,而是利用作者在图书馆长期从事情报检索研究和实践的知识与经验,对各类网络信息检索工具进行深入浅出地分析和介绍,帮助读者系统掌握网络信息检索原理与检索技巧,深入了解各种检索工具的适用领域和使用技巧,使读者成为网络信息检索的驾驭者,在日新月异的网络信息环境中立于不败之地。

“工欲善其事,必先利其器”。为此,本书对网络信息检索工具的基本原理进行了精练介绍,对其检索功能进行了详细分析,帮助读者“知其然,也知其所以然”。本书重点介绍检索工具和数据库的检索方法和技巧,而且,还注明了哪些资源属于免费获取,哪

些资源属于收费获取,这些资源可通过什么途径获得,特别是外文全文数据库的获取途径。

本书内容全面系统、及时新颖。第一章和第二章是 Internet 的基础知识;第三章和第四章介绍了中文搜索引擎和外文搜索引擎的使用方法;第五章和第六章详细介绍了中文数据库和外文数据库检索方法与技巧。

本书适合一切对 Internet 感兴趣的读者及从事信息检索、信息管理的专业人员和教学科研人员。

本书虽经编者们的努力,但由于我们水平有限,撰写时间较为仓促,疏漏、不足乃至错误之处在所难免,真诚地期望得到读者和专家的批评指正。

目 录

第一章 Internet 与数字信息资源

第一节 Internet 概况	3
一、Internet 的历史与发展	5
二、Internet 在中国的发展	8
三、Internet 的基本结构	11
四、Internet 的接入方式	13
第二节 万维网(WWW)数字信息资源	22
一、万维网的结构与特点	22
二、万维网上的数字信息资源	23
三、万维网浏览器	24
第三节 其他 Internet 数字信息资源	27
一、FTP 数字信息资源	27
二、BT 信息资源	27

第二章 搜索引擎

第一节 网络搜索引擎构成及原理	30
一、搜索引擎的定义	30
二、搜索引擎的构成	30
三、搜索引擎的基本原理	30
四、搜索引擎的数据采集机制	31
五、搜索引擎的网页标引机制	32

六、搜索引擎的数据组织机制和用户检索机制	33
第二节 搜索引擎的类型	33
第三节 搜索引擎的检索方法与检索技术	40
一、一般检索步骤	40
二、搜索引擎的检索功能	41
三、智能搜索引擎	41
四、如何选择搜索引擎	44
五、搜索引擎的使用技巧	45
六、优化搜索引擎的方法	48
七、未来的搜索引擎的发展趋势	52
第三章 中文常用搜索引擎的检索方法和技巧	
第一节 Google 搜索引擎	55
第二节 百度搜索引擎	64
一、百度简介及特色功能	64
二、网页搜索	65
三、图片搜索	65
四、MP3 搜索	67
五、资讯	67
六、贴吧	68
七、高级搜索、地区搜索和个性设置	68
八、百度搜索技巧	69
第三节 雅虎搜索引擎	69
一、简介及特色	69
二、网页搜索	70
三、图片搜索的特殊功能	70
四、新闻搜索的特殊功能	71

第四节 搜狗搜索引擎	72
第五节 其他搜索引擎简介	75
第六节 搜索知识与技巧集锦	76
第七节 快速高效搜索的注意事项	78

第四章 外文资料查询网站

第一节 AltaVista	80
一、AltaVista 搜索引擎简介	80
二、设置	82
三、翻译器	83
第二节 Ask/Teoma	83
第三节 Links2go/Go Guides	91
第四节 Openfind	92
第五节 其他一些外文资料查询网站的简介	95
一、Looksmart 搜索引擎简介及登录	95
二、Lycos 搜索引擎简介及登录	95
三、Alltheweb 搜索引擎简介及登录	97
四、Inktomi/msn 搜索引擎简介	99
五、Overture 搜索引擎	100
六、Dmoz/ODP 分类目录简介	100
七、英文搜索引擎控制关系	104

第五章 常用中文数据库

第一节 维普 - 中文科技期刊数据库	106
第二节 中国期刊网与中国期刊全文数据库	118
第三节 万方数据资源系统	130

一、概况	130
二、数字化期刊子系统	131
三、中国学位论文全文数据库	133
四、中国学术会议论文全文数据库	139
第四节 中国生物医学文献数据库	142
一、数据库简介	142
二、数据库检索系统功能介绍	143
三、检索途径与方法	144
四、检索结果的显示和输出	150
第六章 常用外文数据库	
第一节 引文索引	152
一、引文索引概述	152
二、引文检索概述	153
三、美国《科学引文索引》	153
四、ISI Web of Knowledge 平台的主要数据库	154
五、Web of Science 数据库检索	156
六、ISI Web of Knowledge 跨库检索(Cross Search)	171
第二节 工程索引(EI)	174
一、EI 数据库介绍	174
二、检索方法	176
三、检索技术	180
四、检索结果的显示及处理	182
第三节 《科学技术会议录索引》和《社会科学与人文学会议录索引》	182
一、ISI Proceedings 概况	182
二、检索功能	183

三、输出检索结果	185
第四节 OCLC First Search 系统及其检索	185
一、数据库概况	185
二、OCLC FirstSearch 检索方法	188
三、OCLC FirstSearch 检索技术	192
四、检索结果	195
第五节 MEDLINE 和 PubMed 数据库	196
一、美国《医学数据库》(Medline)	196
二、PubMed 数据库	217
第六节 Elsevier Sciences(SDOS)全文数据库	234
一、数据库介绍	234
二、检索方法	234
三、检索技术	238
四、检索结果的显示及处理	239
第七节 Springer - Link 数据库	240
一、Springer - Link 概况	240
二、检索方法	241
三、检索技术	245
四、检索结果的处理	247
第八节 Kluwer online Journals 全文数据库	248
一、概况	248
二、检索方法	249
三、检索技术	253
四、检索结果显示、标记、下载、打印	254

第一章 Internet 与数字信息资源

Internet 有极为丰富的信息资源,而且多数是免费的。虽然 Internet 最初的宗旨是为大学和科研单位服务,但目前它已经成为服务于全社会的通用信息网络。

各种服务器的丰富的信息服务功能和友好的用户接口使 Internet 可以做到雅俗共赏。后面即将看到,Internet 的丰富信息服务方式使之成为功能最强的信息网络,传统网络的功能均包括在内,此外还有许多新的功能。除了 TCP/IP 协议所提供的基本功能外,还有许多高级的信息服务方式和友好的用户接口。以 Gopher 客户程序为例,它可以使用 Internet 上所有 Gopher 服务程序所存储的信息(通常称为 Gopher 空间),而且主机地址和存取路径对用户是完全透明的。Archie, WAIS 和 WWW 的情况也十分类似。这种强大的网络信息服务手段是其他网络难以比拟的。在 WWW 通信协议的基础上开发的 Netscape 浏览器软件是目前水平最高的网络化用户接口,其强大的“导航”功能可以帮助用户在 Internet 的信息海洋中随意漫游。

信息网络的安全问题是当今网络技术的一个重要研究课题。它不仅是一个技术问题,也是一个社会问题和法律问题。要解决信息网络的安全问题,必须采取技术和立法等多种手段进行综合治理。

国外专家在谈到 Internet 时把它称为一个“没有法律,没有警察,没有国界和没有总经理的全球性网络”,也有人把 Internet 称为由用户自己进行管理的“自治网络”。

正是 Internet 的这种开放性,公开性和自治性,使网络在安全方面难以尽如人意。1988 年 11 月 2 日晚,Internet 首次受到了“蠕虫式病毒”的严重侵袭,使数千台计算机无法正常工作。所谓“蠕

虫式病毒”实际上是一种自我复制程序,它利用了某些网络程序的弱点,使大量的上网主机受到感染。

总而言之,使用 Internet 的用户要了解一点,那就是在 Internet 的环境下数据的安全性应该由数据的拥有者和维护者自己负责。

当然,除了安全性方面的缺点之外,Internet 也还有一些其他美中不足之处。例如,由于信息资源的分散化存储和管理,给用户在查找 Internet 资源方面带来一定的困难;种类繁多的服务方式在给用户提供使用的灵活性的同时,也给一些计算机和网络知识比较缺乏的用户造成某种不便;自由化的发展模式在赢得广大用户欢欣的同时也使一些不宜广泛传播的信息(如黄色照片和录像等)失去控制。所有这些都体现了所有现代化科学技术的两面性,需要我们用唯物主义的态度和批判的眼光有选择地去接受它们。

Internet 上的信息资源可分为付费和免费两大类。付费的网络资源大多是技术含量高,整理有序,具有很高利用价值的各信息公司开发的数据库系统。如:DIALOG、BRS、ORBIT 等等。Internet 上的数据库不计其数,除了 DIALOG、STN 所提供的收费商业性数据库外,也有许多免费供用户使用的数据库。除文献数据库外,还有各种如招聘求职库、征婚交友库、公司名录库、专利库、标准库等,涉及各个领域、不同专业。例如美国、加拿大、日本等国的专利数据库在 Internet 上均可得到免费使用。免费信息资源大多是由单位或个人提供的各种信息资源。

Internet 的主要信息服务分为以下几类:

- 电子邮件 E-Mail
- 文件传输 FTP
- 远程登录 Telnet
- 电子公告板 BBS
- 信息浏览 Gopher
- 高级超文本浏览 WWW

- 自动标题搜索 Archie
- 自动搜索 WAIS
- 域名系统 DNS

这些服务中,常用的是电子邮件、高级超文本浏览和文件传输。其中的自动标题搜索和自动搜索由于搜索命中率不高或使用效率不高而逐渐被 WWW 的新的多功能高效率的搜索工作如 yahoo 等所代替。DNS 是 Internet 上所固有的功能性服务,是为其他服务而提供服务的。

第一节 Internet 概况

Internet 是一个建立在 TCP/IP 协议族上的国际互连网络,它是各个子网以网状结构互联而成,在每个子网中存在着数量不等的主机,子网及其主机均以 IP 协议统一编址。子网可以是 LAN,也可以是 WAN。主机可以是网上的客户机、服务器或者路由器等设备。如图 1-1 所示。

一般说来,可以这样描述 Internet:Internet 是一个遵从 TCP/IP 协议,将大大小小的计算机网络互联起来的计算机网络。Internet 不仅连接了世界各地的数千万台计算机,更重要的是把计算机前的所有人联系在一起,形成在 Internet 上工作、学习和生活的一个社会网络,创造了一种前所未有的人文环境。在它下面连接地区性网络,地区性网络与广域网(WAN)相连接,广域网连接局域网(LAN),局域网里连接着许多计算机。这样,把许多计算机连接在一起,实现资源共享。本世纪初,Internet 将连接近亿台计算机,达到以十亿计的用户,Internet 对个人意味着什么?对于终端用户说来,Internet 不是指一个网络而是若干网络的集合这一点,不具有什么重要的意义。你首要考虑的是作某些事情,例如运行一个程序,访问需要的数据,或者交换有用信息。这种时候,你不必关心 Internet 是如何组合在一起的,或者被访问对象是在哪里。

Internet 是成千上万信息资源的总称,这些资源在线的分布在

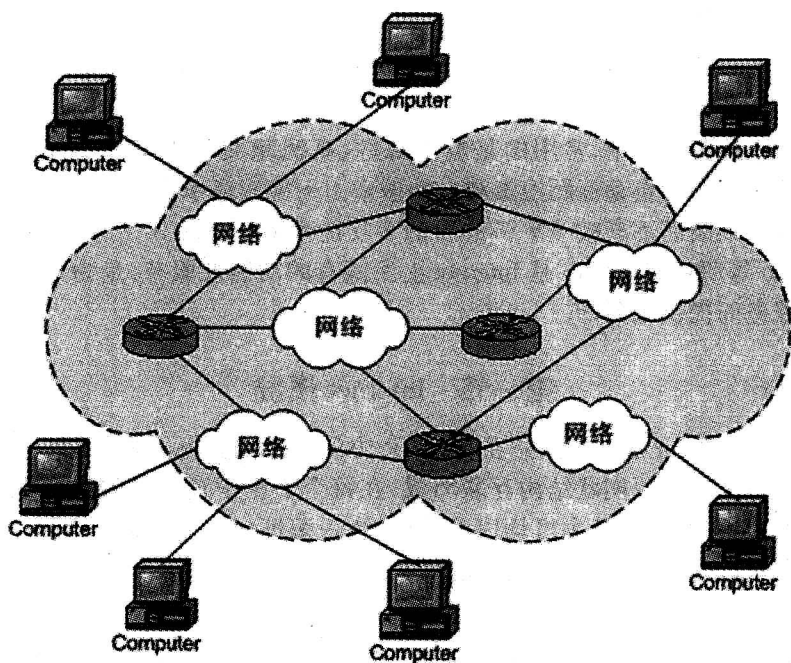


图 1-1 Internet 网络结构

世界各地的数百万台计算机上。Internet 是一个社会大家庭,家庭成员可以方便的交换信息,共享资源,Internet 上开发了许多应用系统,供网上的用户使用。Internet 目前所推行的发展战略与许多国家实施信息高速公路计划所追求的“全民服务”的目标更是一脉相承。因此,Internet 不仅在计算机网络发展史上占有重要地位,而且将作为信息高速公路的基础,不断向前发展。自 94 年以来,Internet 开始了商业化的发展,利用 Internet 进行商业活动成为世界经济的一大热点,几乎所有的国际著名公司都着手在 Internet 上建立自己的商业服务系统,并且把公司管理系统与 Internet 相连接。另一方面,商业性 Internet 接入服务也为其带来了更多的用户,推动了 Internet 的普及,甚至对人们的生活方式产生了重大

影响。可以说 Internet 的普及应用,是人类社会由工业社会向信息社会发展的重要标志。

一、Internet 的历史与发展

举世瞩目的 Internet 是由美国 20 世纪 60 年代的 Arpanet 网络发展和演化而成的。Internet 的形成与发展,经历了试验研究网络、学术性网络,以及商业化网络这 3 个历史阶段。

1. 试验研究网络 1969 年,美国国防部的国防高级研究计划署(Advanced Research Project Agency,简称 ARPA)建立了一个采用存储转发方式的分组交换广域网——Arpanet,该网络仅有 4 个节点,分别建在加州大学洛杉矶分校(UCLA)、斯坦福研究所(SRI)、加州大学圣大巴比分校(UCSB),以及犹他大学(UTAH),该网络是为了验证远程分组交换网的可行性而进行的一项试验工程,以防止核战争爆发引起大量电话业务中断导致军事通信瘫痪的局面出现。Arpanet 就是今天 Internet 的前身。

虽然初期 Arpanet 各节点之间的连接只能使用 56Kbps 的专线,但是网络的扩展是相当惊人的。1972 年首届国际计算机通信会议(IGCC)上首次公开展示了 Arpanet 的远程分组交换技术,当时 Arpanet 已有约 20 个分组交换节点机和 50 台主机。在总结最初的建网实践经验的基础上,开始了被称为网络控制协议(NCP)的第二代网络协议的设计,ARPA 随后又组织有关专家开发了第三代网络协议——TCP/IP 协议,该协议于 20 世纪 70 年代中期由斯坦福大学的 Vinton Cerf 和 BBN 的 Robert Kahn 开发,1983 年正式在 Arpanet 上启用,这是 Internet 发展史上的一个里程碑。1983 年,ARPA 将网络控制权交给防卫通信局(DCA),并将 Arpanet 分割成两个部分,一部分是专用于国防的 MilNET(一个非保密的军事通信网络),另外一部分仍然称为 Arpanet。与此同时,美国还相继建成了 CSNET 和 BITNET 两个网络。Arpanet 的建立,产生了网络互联的概念。