



网海拾珠

WWW资源的查询

主编 张 奇 陈顺忠

上海科学技术文献出版社

网 海 拾 珠

——WWW 资源的查询

主 编：张 奇 陈顺忠

编 委：陈 喆 彭 伟 赵文绮 钮 佳

图片剪辑：赵文绮

插图设计：钮 佳

上海科学技术文献出版社

图书在版编目(CIP)数据

网海拾珠:WWW 资源的查询/张奇,陈顺忠编著.
上海:上海科学技术文献出版社,2002.7
ISBN 7-5439-1990-7

I. 网... II. ①张...②陈... III. 万维网-情报检
索 IV. G252.7

中国版本图书馆 CIP 数据核字(2002)第 038799 号

责任编辑:曹文青

封面设计:石亦义

网 海 拾 珠

——WWW 资源的查询

主编 张 奇 陈顺忠

*

上海科学技术文献出版社出版发行

(上海市武康路 2 号 邮政编码 200031)

全国新华书店经销

江苏常熟人民印刷厂印刷

*

开本 850×1168 1/32 印张 7.75 字数 180 000

2002 年 7 月第 1 版 2002 年 7 月第 1 次印刷

印数:1-3 100

ISBN7-5439-1990-7/T·697

定 价:16.80 元

前 言

随着互联网(Internet)成功商业化,它已成为世界上规模最大、用户最多、影响最广的网络互联系统。上网冲浪已成为互联网弄潮儿获得信息的主要手段之一。互联网资源的佼佼者 WWW 资源的发展对互联网的前进更是起了推波助澜的作用。同时互联网也是一把双刃剑,一方面它推动了数字技术和信息技术的发展,另一方面却带来了大量的垃圾信息,困扰人们的视觉,同时又给不谙上网之道的人造成了数字鸿沟。

本书以帮助人们消除数字鸿沟,滤清信息垃圾为宗旨,采用深入浅出的写作手法和图文并茂的风格从互联网基础知识开始,带您走上采集信息之旅。工具篇中送给您一套信息搜索的工具,并通过资源查询实例让您有身临其境的感受,从中了解获取信息的最佳方法,这样您就可以轻松地不会到会了。本书的资源篇和提高篇旨在让已经是网海中的徜徉者获得一些理论知识、辅助工具和资源积累,在 WWW 资源查询中更得心应手。衷心希望您能通过本书成为 WWW 资源查询的高手,从浩瀚的信息海洋中拾得无价

之宝。

本书是在上海图书馆上海科技情报研究所网络服务中心的几位同仁的共同努力下完成的。基础篇由赵文绮、张奇撰写；工具篇由陈顺忠、张奇、陈喆、彭伟撰写；方法篇由陈喆、陈顺忠、张奇撰写；资源篇由陈喆、钮佳、赵文绮撰写；提高篇由张奇、彭伟撰写。最后，张奇、陈顺忠对全书进行统稿和审校。

应了“不是我不明白，而是这世界变化快”的歌词，我们在编撰本书的过程中，发现许多网站和搜索引擎的功能以三天大变样的速度变化着，给编撰带来了一定难度，因此需要声明的是，除了文中的特别脚注，本书文字内容和图片内容均以交稿日期2002年4月28日为准。由于编撰者知识范围所限，书稿中难免有疏漏和欠妥之处，敬请指正。

张 奇 陈顺忠

第一章 基础篇	3
一、什么是 Internet?	3
二、Internet 上有哪些资源?	4
三、什么是 WWW?	4
四、什么是浏览器 (Browser)?	6
五、您会使用 IE 吗?	7
1. 如何浏览 WWW 资源?	8
2. 如何让浏览页面前进和后退?	8
3. 如何收藏我喜爱的网页?	9
4. 如何将喜爱的网页收藏到指定的目录中?	9
5. 如何备份收藏夹?	10
6. 为什么我的 IE 窗口没有地址栏?	11
7. 如何设置主页?	11
8. 如何快速打开新的浏览窗口?	12
9. 我可以将网页保存下来吗?	13
10. 如何与我的好朋友分享一个精彩的网页?	14
11. 如何查看最近访问过的网页?	14
12. 我的 IE 出现不正常现象,有什么方法可以修复吗?	15
13. 我可以随意改变显示文字的大小吗?	16
14. 我想保护自己的上网隐私,怎么办呢?	17

CONTENT

目 录

六、什么是 Usenet?	18
七、什么是 Mailing List?	19
八、什么是 WWW 资源搜索工具?	20
1. 什么是搜索引擎?	20
2. 什么是目录搜索引擎?	21
3. 什么是元搜索引擎?	21
九、常用的搜索方法有哪些?	22
1. 什么是分类检索?	22
2. 什么是关键词检索?	23
3. 什么是词组检索?	23
4. 什么是布尔逻辑检索?	23
5. 什么是加或者减号检索?	24
6. 什么是截词检索?	24
7. 什么是字段检索?	25
8. 什么是自然语言检索?	25
9. 什么是位置检索?	25
10. 什么是概念检索?	26
 第二章 工具篇	 29
一、搜索引擎	29

1. Google	29
2. LYCOS 中国	41
3. 新浪搜索引擎	48
4. 雅虎中国(Yahoo!)	53
5. 搜狐(Sohu)	58
二、元搜索引擎	62
1. DOGPIL	62
2. ProFusion	74
3. Ixquick	84
三、专业搜索	90
1. 股票行情搜索——证券之星	90
2. 图像搜索——VisionNEXT	95
3. 地图搜索——图行天下	97
4. MP3 搜索——搜刮网	103
5. 书目查询——WebPAC	107
6. 图书搜索——当当书店	111
7. 期刊论文搜索——中国期刊网	113
8. 标准搜索——中国标准咨询网	120
9. 企业产品搜索——阿里巴巴	123
10. 法律搜索——北大法律信息网	128
11. 医药搜索——37℃医学网	132

12. Mailing List 搜索——希网网络	137
第三章 方法篇	143
实例一	144
实例二	153
实例三	161
实例四	164
第四章 资源篇	171
一、搜索引擎特性表	171
1. About	171
2. AOL	172
3. Excite	173
4. Hotbot	174
5. Infoseek	175
6. Lycos	176
7. MSN	177
8. Northernlight	178
9. Overture	179
10. WebCrawler	180

11. 百度	181
12. 北大天网	182
13. 北极星	183
14. 焦点	184
15. FM365	185
16. 网络指南针	186
二、元搜索引擎特性表	187
1. Highway61	187
2. Mamma	188
3. One2seek	189
4. Savvysearch	190
5. Surfwax	191
6. Vivissimo	192
三、资源列表	193
1. 百科全书	193
2. 财经	193
3. 参考咨询	193
4. 词典	194
5. 出国留学	194
6. 地图	194
7. 电脑网络	195

CONTENT

目 录

8. 电子图书	195
9. 法律	196
10. 房产	196
11. 黄页	196
12. 教育培训	197
13. 旅行	197
14. 期刊论文	198
15. 求职	198
16. 生活服务	199
17. 书店	199
18. 图书馆书目查询	200
19. 网上购物	200
20. 网络文学	201
21. 医药	201
22. 音乐	202

第五章 提高篇 205

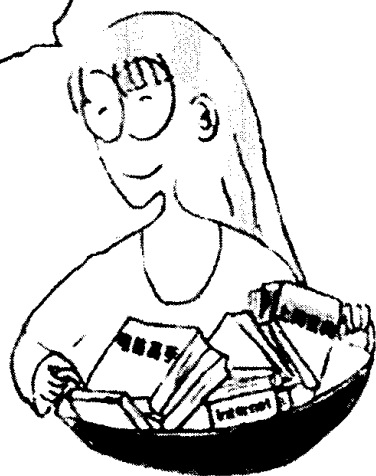
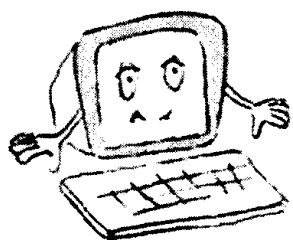
一、原理	205
1. 搜索引擎是如何工作的?	205
2. 目录搜索引擎的原理是什么?	207

CONTENT

目 录

3. 目录索引与搜索引擎有什么区别?	208
4. 元搜索引擎是如何工作的?	209
5. 元搜索引擎有哪些种类?	210
6. 如何评价和选择元搜索引擎?	212
二、辅助工具	213
1. 离线浏览——Teleport Pro	213
2. 断点续传工具——FlashGet	220
3. 压缩解压工具——WinRAR	227
4. 图浏览工具——ACDSee	232

咱们从基础
做起吧!



第一章 基础篇

一、什么是 Internet?

Internet,称为国际互联网,或因特网。它是由遵循 TCP/IP〔1〕协议的世界上成千上万个计算机网络组成的全球网络。

说起 Internet,我们不得不提起美国国防部于 1969 年建立的计算机互联网络 ARPANet〔2〕。它是互联网的鼻祖。由于 ARPANet 仅供军事部门和少数科学家使用,为了教育与科研的需要,美国科学基金会〔3〕于 1984 年出资建立了 NSFNet。由于 NSFNet 的成功,它建成后取代了 ARPANet 而成为 Internet 的主干网。

随着 Internet 成功商业化,它已成为世界上规模最大、用户最多、影响最广的网络互联系统。据 2002 年 3 月 INT Media Group 以及 2002 年 1 月中国互联网中心(CNNIC)统计数据显示,目前国际互联网上网计算机已达 14 亿台,上网人数达 44.6

〔1〕 TCP/IP(Transmission Control Protocol/Internet Protocol,传输控制协议/互联网协议),为国际互联网主要通讯协议。

〔2〕 ARPANet 全称: Advanced Research Projects Agency Net。

〔3〕 美国科学基金会 National Science Foundation,NSF。

亿人;我国上网计算机已达1 254万台,上网人数达3 370万人,国际出口带宽总量为7 597.5M,国际出口为8个,连接到美国、加拿大、澳大利亚、英国、德国、法国、日本、韩国等国。

二、Internet 上有哪些资源?

Internet 是世界上最大的信息宝库,它覆盖面广,包罗万象。从它所涉及的领域来看,您可以获取社会、政治、历史、教育、卫生、科技、自然、经贸、金融、文化、艺术、娱乐等各个方面的内容信息;从它提供的语言文种来讲,您可以得到中文、英文、法文、日文、俄文等各语种的信息资源;从它所采用的网络传输协议来划分,为您提供了 WWW、Telnet〔1〕、FTP〔2〕、新闻组、邮件列表、讨论组等信息来源;另外,除文字信息以外,还有图表、图形、图像、声音、动画、影像等类型资源。

所以,有人说 Internet 无奇不有,无所不存,您可以在网上找到需要的几乎所有信息,快捷方便地与别人交换信息,访问各个领域的资深专家。不过,由于没有专门机构对 Internet 负责,您还需对查找到的信息辨别一下真伪。另外,有些商用的数据库是需要付费访问的。当然,您也别担心,Internet 的信息资源绝大多数是真实可靠的。

三、什么是 WWW?

WWW 是 World Wide Web 的简称,亦称 Web、万维网或环

〔1〕 Telnet : Telecommunications Network,远程通讯网。

〔2〕 FTP: File Transfer Protocol,文件传输协议。

球网。它由欧洲核子研究中心于 1972 年开发。由于它采用超文本技术和多媒体技术,支持多种网络协议,在同一界面上同时可以使用 FTP、Telnet 等其他 Internet 服务,WWW 推出之后很快成为发展最快和应用最广的互联网信息服务。

自 WWW 问世以来,互联网上拥有 5 000 个以上网页的 Web 服务器已达 38 444 856 台,2002 年 2 月比 1 月增加了 175 万台〔1〕。WWW 资源成为互联网信息资源的主体,没有 WWW 就没有今日 Internet 的辉煌。

所谓 WWW,是建立在客户机/服务器模型之上,以 HTML〔2〕语言和 HTTP〔3〕协议为基础,以统一的图形用户界面提供各种 Internet 协议服务的网络资源与查询工具,其结构图见图 1-1。



图 1-1 WWW 网络传输模型

WWW 服务器上存贮的是机构或个人需在 Internet 上发布的各种类型的信息,包括文本、图形、图像、声音、动画、影像

〔1〕 来自 Netcraft 网站 2002 年 2 月的调研报告。

〔2〕 HTML(Hypertext Markup Language,超文本标记语言),是用于创建 WWW 文档的编程语言。

〔3〕 HTTP(Hypertext Transfer Protocol,超文本传输协议),是 WWW 客户机与服务器之间应用层的通讯协议。