

工具软件实用详解系列

Internet

工具

和翻译软件

TRANSLATE

实用详解

蹇文 周宇坤 编著

清华大学出版社

<http://www.tup.tsinghua.edu.cn>



TP393.4

Q34

453323

Internet 工具和翻译 软件实用详解

骞 文 周宇坤 编著



3



00453323

清华大学出版社

(京)新登字 158 号

JS208 / 22
内 容 简 介

本书是一本 Internet 工具书, 全面介绍了 Internet 用户经常使用到的工具, 包括搜索工具、辅助浏览工具、站点下载工具、文件传输工具、BBS 登录工具、电子邮件工具、MUD 工具等, 同时, 还介绍了经常使用的翻译软件。通过使用本书, 用户可以有效地提高在 Internet 上浏览的效率。本书适合所有 Internet 用户阅读。

版权所有, 翻印必究。

本书封面贴有清华大学出版社激光防伪标签, 无标签者不得销售。

图书在版编目 (CIP) 数据

Internet 工具和翻译软件实用详解/骞文, 周宇坤编著. -北京: 清华大学出版社, 1999.7
ISBN 7-302-03738-8

I . I... II . ①骞... ②周... III . ①因特网-软件工具②因特网-翻译-应用软件
IV . TP393.4

中国版本图书馆 CIP 数据核字 (1999) 第 50134 号

出版者: 清华大学出版社 (北京清华大学校内, 邮编 100084)

<http://www.tup.tsinghua.edu.cn>

印刷者: 北京丰华印刷厂

发行者: 新华书店总店北京发行所

开 本: 787×1092 1/16 印张: 22.5 字数: 557 千字

版 次: 1999 年 11 月第 1 版 1999 年 11 月第 1 次印刷

书 号: ISBN 7-302-03738-8/TP · 2091

印 数: 0001~8000

定 价: 28.00 元

前 言

随着 Internet 在各个领域的广泛发展，其功能越来越多地发挥出来，构成了丰富多彩的网络交流体系。这些功能包括：主页浏览（http 超文本传输），WWW 搜索引擎，文件传输 FTP，电子公告板 BBS，电子邮件 E-mail，网络游戏 MUD 等。正是由于这些功能的开发和完善，才使得整个网络生机勃勃，成为我们生活中不可缺少的东西。

但是，当您渴望成为网络大家庭的一员时，却往往因为不熟悉网络工具，不知道如何操作它们而信心不足。也许您时常会被这样的问题所困扰：

在茫茫的网点中，如何才能找到自己感兴趣的站点？

怎样才能知道需要的软件在哪里？

怎样才能欣赏网络上的多媒体呢？

网页的传输能不能更快些？

该怎样做才能实现离线浏览呢？

FTP 可以带来大量的有用软件，可该怎样进行文件的传输？

BBS 也很吸引人，可怎样才能加入到电子公告牌中？

如何收发电子邮件？

如何申请免费的 E-mail 帐号？

如何进入虚幻而又真实的网络“泥巴”世界？

如何防止某些恶意的家伙破坏自己的网络？

什么样的汉化软件才是最适合自己的？

怎样把汉化软件的作用充分发挥出来？

.....

如果您对这些问题感到困惑不解，那么在看完本书之后，这些疑问也许都会迎刃而解。本书的第 1 章至第 8 章网络常用软件详解将告诉您如何来操作一些常用的网络软件，第 9 章的翻译软件大观则为您选用合适的汉化软件提供依据。当然，这还不仅仅是全部。因为本书始终贯穿着这样一个思想：不是以介绍一个软件为目标，而是以介绍一类软件为目标。所以，读者可以借助这本书来触类旁通。

同时，这本书也可以作为速查手册使用。您可以针对自己所迫切需要解决的问题来有选择地学习其中的部分章节。因为任何一个章节乃至任何一种工具都是独立的。

本书的第 1、2、5、9 章由周宇坤完成；第 3、4、6、7、8 章由骞文完成；其余部分则由双方共同搜集资料完成。尽管本书在写作过程中参考了包括《计算机世界》、《Internet 宝典》在内的权威性报刊及书籍；必要时还上网验证某些结论，但毕竟网络信息瞬息万变，书中如存在不当之处或者叙述欠缺的地方，还请读者批评指正。

作者

1999 年 5 月

目 录

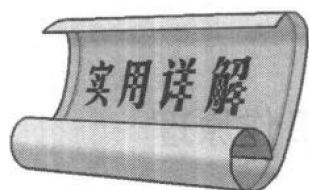
前言.....	I
第 1 章 WWW 上的搜索引擎.....	1
1.1 搜索引擎的主要作用及其工作原理.....	2
1.2 优秀搜索引擎的特点.....	3
1.3 搜索引擎语法规则.....	5
1.4 搜索策略与技巧.....	7
1.5 中文搜索引擎介绍.....	8
1.6 国外搜索引擎.....	23
第 2 章 WWW 浏览辅助工具.....	31
2.1 代理猎手(Proxyhunter).....	32
2.1.1 “代理猎手”到底能干什么.....	32
2.1.2 “代理猎手”界面介绍.....	32
2.1.3 启用“代理猎手”.....	33
2.1.4 如何在浏览器中设置代理服务.....	36
2.1.5 “代理猎手”的内部参数设置.....	38
2.2 网络超音鼠(NetSonic).....	39
2.2.1 “网络超音鼠”是怎样使网络加速的.....	39
2.2.2 启动“网络超音鼠”.....	40
2.2.3 查看帮助技巧.....	40
2.2.4 定制自己的“超音鼠”——参数设置问题.....	41
2.2.5 禁用和启用“超音鼠”.....	42
2.2.6 图标闪烁问题.....	43
2.3 网络媒体播放器(RealPlayer).....	43
2.3.1 网络发展进入多媒体阶段.....	43
2.3.2 RealPlayer 的 3 种界面.....	43
2.3.3 RealPlayer 的使用步骤.....	45
2.3.4 RealPlayer 的快捷频道.....	46
2.3.5 RealPlayer 的参数设置.....	46

第 3 章 站点下载工具	53
3.1 为什么要离线浏览	54
3.2 Teleport Pro——功能繁多的站点捕捉者	54
3.2.1 关于 Teleport Pro	54
3.2.2 项目控制	55
3.2.3 运行控制	61
3.2.4 项目属性	63
3.2.5 自动连接和定时服务	70
3.2.6 使用命令行	73
3.2.7 设置代理服务器	73
3.3 BlackWidow (黑寡妇)的手段	74
3.3.1 BlackWidow 简介	74
3.3.2 “黑寡妇”主窗体介绍	74
3.3.3 快速开始	75
3.3.4 菜单命令	79
第 4 章 文件传输工具——FTP	83
4.1 FTP 的由来	84
4.2 WS_FTP Pro——从 Win 3.1 时代就和我们相识的老朋友	84
4.2.1 WS_FTP Pro 的初步登录以及使用	84
4.2.2 WS_FTP Pro 的设置	93
4.2.3 WS_FTP Pro 站点管理器(WS_FTP Pro Explorer)	99
4.3 CuteFTP——具备断点续传功能的朋友	101
4.3.1 CuteFTP 的快速开始和主界面	102
4.3.2 CuteFTP 的选项设置和高级功能	109
4.4 LeapFTP——一个好用的 FTP 客户软件	115
4.4.1 LeapFTP 的综述	115
4.4.2 LeapFTP 的脚本语言提要	118
4.5 GetRight——附着在浏览器上的文件传输器	119
4.5.1 GetRight 综述	120
4.5.2 让 GetRight 工作	120
4.5.3 GetRight 的工作控制	120
4.6 Mr.Cool	129
4.6.1 使用 Mr.Cool	130
4.6.2 Mr.Cool 的设置	131
4.7 NetAnts	135
4.7.1 网络蚂蚁的设置	135
4.7.2 工作实例	137

4.8 本章小结	139
第 5 章 BBS 登录工具	141
5.1 关于电子公告牌	142
5.2 NetTerm——最常用的 BBS 登录工具	142
5.2.1 NetTerm 的功能概要	142
5.2.2 NetTerm 的界面介绍	143
5.2.3 启动远程登录	143
5.2.4 屏幕字体、颜色及桌面的设置	145
5.2.5 选项	148
5.2.6 屏幕拷贝与跟踪	151
5.3 Tera Term	152
5.3.1 Tera Term 的特点	152
5.3.2 开始新连接	152
5.3.3 File 菜单的其他功能	153
5.3.4 Edit 菜单	155
5.3.5 参数设置	155
5.3.6 终端控制	159
5.4 Cterm——聊天功能最完善的 BBS 登录工具	159
5.4.1 Cterm 的特色	159
5.4.2 Cterm 的地址簿管理及启动连接	160
5.4.3 调整 BBS 显示到最佳状态	161
5.4.4 Cterm 的鼠标操作技巧	162
5.4.5 自动全文拷贝	162
5.4.6 聊天室高级技巧	163
5.4.7 封锁与解封锁 BBS	165
5.4.8 离站	166
5.5 Telnet	166
5.5.1 Telnet 启动连接和断开连接	166
5.5.2 Telnet 的全部复制功能	167
5.5.3 终端设置	167
5.5.4 开始记录和停止记录	168
第 6 章 E-mail 工具	171
6.1 Eudora——经典的 E-mail 软件	172
6.1.1 Eudora Pro 的主界面	172
6.1.2 设置	172
6.1.3 接收和发送信件	181
6.1.4 管理邮件	184

6.1.5	信纸模板(Stationery).....	187
6.1.6	地址簿.....	188
6.1.7	过滤器.....	190
6.2	Fymail 飞扬电子邮件——人性化界面的中文 E-mail 软件.....	192
6.2.1	Fymail 的服务器和帐号设置.....	193
6.2.2	飞扬电子邮件的人性化界面.....	194
6.2.3	信件处理和远程邮件管理.....	199
6.2.4	通讯录管理.....	203
6.2.5	信件加密.....	204
6.3	FoxMail.....	205
6.3.1	FoxMail 的传统界面.....	205
6.3.2	设置服务器、帐户以及其他选项.....	206
6.3.3	邮件处理.....	209
6.3.4	地址簿管理.....	212
6.3.5	信箱管理和邮箱助理.....	213
6.3.6	远程邮箱管理和邮箱监视器.....	214
6.4	免费电子邮件.....	216
6.4.1	“永动邮箱”综述.....	216
6.4.2	电子邮箱的申请.....	217
6.4.3	登录进入免费邮箱.....	221
第 7 章	MUD 工具——zMUD.....	229
7.1	zMUD 的使用初步.....	230
7.1.1	连接 zMUD 站点, 快速建立身份.....	230
7.1.2	zMUD 的参数设置.....	231
7.2	zMUD 的高级使用.....	244
7.3	地图编辑器(Mapper).....	249
第 8 章	网络安全.....	255
8.1	邮件安全的介绍和防范.....	256
8.2	一些网络攻击工具和防范工具的介绍.....	258
8.3	Win95/NT 的 BUG 及如何修正这些 BUG.....	261
第 9 章	翻译软件大观.....	263
9.1	引言.....	264
9.1.1	博览群雄——当今汉化软件发展趋势.....	264
9.1.2	翻译软件排行榜.....	268
9.2	东方快车 2000.....	268
9.2.1	东方快车的基本工具条界面.....	269
9.2.2	高级按钮.....	275

9.2.3 东方快车全文翻译系统	277
9.3 即时汉化专家 98 增强版	281
9.3.1 即时汉化专家的工具条界面和功能概述	281
9.3.2 系统设置	284
9.3.3 集成环境	285
9.3.4 词库集成系统	288
9.4 通译 Internet 2.0 在线翻译系统	291
9.4.1 关于通译在线翻译系统	291
9.4.2 通译在线翻译	293
9.4.3 电子词典系统	296
9.4.4 全文翻译系统	300
9.5 汉神网际套餐	302
9.5.1 汉神网际视窗	302
9.5.2 汉神网际快译	307
9.5.3 汉神电子辞典	309
9.6 网际金典 ROBOWORD3.0	311
9.6.1 网际金典概述	311
9.6.2 运行网际金典	312
9.6.3 网际金典设置	314
9.6.4 字典管理	318
9.6.5 学习本	322
9.7 RichWin97 专业版	323
9.7.1 RichWin97 专业版简介	323
9.7.2 RichWin97 中文平台功能详述	324
9.8 金山词霸III	331
9.8.1 金山词霸III概述	331
9.8.2 金山词霸的即指即译	331
9.8.3 词典查询	332
9.8.4 系统设置	334
9.9 地球村 2.0PLUS	338
9.9.1 地球村的主界面	338
9.9.2 提示条	341
9.9.3 词典配置	342
9.10 南极星 2.0(NJWIN COMMUNICATOR)	345
9.10.1 南极星简介	345
9.10.2 南极星的动态工具条界面	346
9.10.3 动态按钮条	346



第 1 章 WWW 上的搜索引擎

搜索引擎的主要作用及其工作原理
优秀搜索引擎的特点
搜索引擎语法规则
搜索策略与技巧
中文搜索引擎介绍
国外搜索引擎

随着 Internet 的日益发展, Internet 上的信息也正在以爆炸性的速度不断增加。以 Yahoo 的统计数字为例, 每天新增的网址以 10^3 的速度增加着。不得不承认, 无论是对于网络冲浪的朋友还是对于借助网上进行商务活动的用户, 信息越丰富就意味着机会越多, 可得到的收益也越大。但这只是表面现象。待熟悉了网络, 就容易发现: 信息分布是无规律的。在 Internet 网上网站林立, 搜集者要想访问到所有自己感兴趣的站点, 对所有相关信息进行一番考察, 恐怕是不可能的。若不借助其他工具, 每个人能访问到的区域实在是微不足道的一小部分。因此和网络打交道, 面临的一个突出问题是: 如何在上百万个网站中快速有效地找到自己想要的信息。

这正是 WWW 搜索引擎(Search Engine)要解决的问题。它们分布在网络上, 就像茫茫大海中的导航灯一样, 针对访问者的要求为其指明该往何处航行。通过搜索引擎, 网络上零零碎碎的网站相对集中、分门别类地会聚到访问者面前, 使访问者可以轻易地寻访到自己需要的站点。

1.1 搜索引擎的主要作用及其工作原理

“引擎”的原意是“发动机”的意思。这意味着搜索引擎最根本的任务是为网络访问者提供一个高效访问的机制。其实, 搜索引擎本身也是个网站, 但这个网站和一般网站不同的是: 一般网站的数据库中存放的是自己的信息, 而搜索引擎的数据库中存放的是其他网站的信息; 一般网站仅仅是接受用户的访问, 在有请求的时候向用户发送自己的网页, 而搜索引擎不仅要接受用户的访问, 还必须根据用户的搜索要求——通常称为关键字(Keyword)——在庞大的数据库中搜索出符合条件的站点, 再返回给用户。因此, 搜索引擎的背后都必须有一套高效的搜索程序, 以支撑起整个引擎。

对于各种搜索引擎, 其工作机制基本相同。通常包括下面所述的 3 个过程。

1. 派出“网页搜索程序”在网络上搜寻新的网页信息, 并将这些信息带回搜索引擎

“网页搜索程序”是搜索引擎的左膀右臂。搜索引擎之所以能够知晓那么多的网站, 并不是因为它总能从负责站点管理的机构那里得到新上网的站点地址, 而是由于搜索引擎每天都会派出绰号为“蜘蛛(Spider)”或“机器人(Robot)”的网页搜索软件, 在各网址上爬行。它们将访问网络中公开区域的每一个站点并记录其网址。如果出现新增的网站, 软件就会创建一个详尽的网络目录。这里面包括了新网站的各种信息, 如网页的数量、网页的层次、网页中的内容等。不同的搜索程序在记录信息时会有些差别, 但是, 其基本目的都是为了获取新网站或者旧网站更新后的第一手信息。

2. 将信息进行分类整理, 建立搜索引擎数据库

在信息的分类整理阶段, 不同的系统会在搜索结果的数量上和质量上产生明显的不同。同时在对信息进行归类的工作中也会依据不同的分类准则。有的系统是把“网页搜索程序”发送往每一个站点, 记录下每一页的所有文本的内容; 而有的系统则首先分析数据库中的地址, 通过测定该站点的链接数量以评估哪些站点最受欢迎, 然后再利用软件记录

这些站点的信息，包括标题、文本内容和摘要。对于任何一种方式而言，都是一件繁琐的工作。然而，搜索引擎不得不这样做，因为搜索引擎必须保证数据库的内容及时地更新，保持与整个世界的信息发展同步。

3. 通过网络服务器端软件，为用户提供浏览器界面下的信息查询服务

访问者是通过 WWW 来访问搜索引擎的。因此，搜索引擎必须具有一个良好、实用的界面。用户可以方便地输入自己要搜索的关键字或者短语，然后递交给搜索引擎，搜索引擎要尽可能快而准确地从数据库中找到符合要求的结果，并把此结果以超链接的形式反馈给用户。用户只要点击相应的超链接，就可以链接到所要去的网站。

优秀的搜索引擎为了实现快速索引，在分类整理阶段就把站点进行了归类，用户可以在各种不同类别中进行搜索。这样不仅提高了搜索的“命中率”，而且也节约了搜索的时间。

该过程表述如图 1.1 所示。

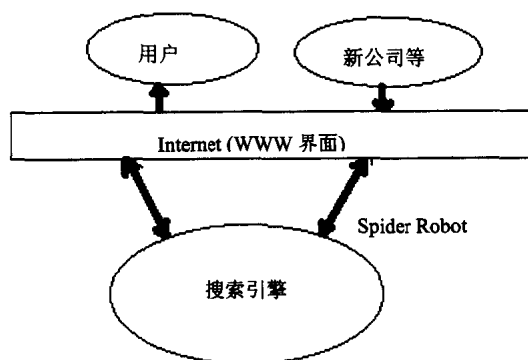


图1.1 搜索引擎工作原理

1.2 优秀搜索引擎的特点

在目前的 Internet 上有各式各样的搜索引擎。国外较著名的有 Yahoo、InfoSeek、AltaVista、Powersearch 等，而国内较著名的则有“网络指南针”、中文 Yahoo、搜狐 Sohu、“悠游”、“天网”、“常春藤”等。单论引擎的建设，显然是国外搜索引擎占优，但是从引擎的适用性来看，如果需要寻找国内站点，还是国内的搜索引擎用起来比较顺手。而一个优秀的搜索引擎，通常都应该具有以下特点。

1. 丰富的索引数据库

这是确保用户能找到自己想要的信息的基本保证。不管怎么说，让用户几分钟后面以一纸空文是很不友好的事情。哪怕搜索到的站点很小，很不起眼，也要比对用户说“抱歉”来得好。否则，沮丧之余，用户很可能不再来访问。

2. 支持文本全文检索或软件全名/缩略名检索

对于文本网页,进行搜索时应当是针对全部文本进行的。这意味着用户可以搜索每一个网页中的每一个词。只要用户给出的关键字在网页中出现过,那么这篇网页就将作为匹配的结果出现在检索结果中。至于提供下载的软件搜索引擎,因在服务器上存储的软件有可能使用缩略名(比如文本编辑器 Ultraedit,其缩略名为 Uedit),而实际记忆中用户往往记的是全名(Ultraedit),因此软件搜索引擎不能是仅仅支持搜索软件缩略名,同时应支持对软件的全名进行搜索。

3. 支持目录式分类结构

搜索引擎最艰巨的工作在于对信息的分类整理。一个具有分类功能的搜索引擎可以保证用户在搜索时获得较好的结果。因为分类搜索引擎在遇到一个网站的时候,并不是简单地把文本内容抓录下来,而是首先将该网站划归到某一个分类下,再记录一些摘要信息,建立起关于这个网站的简要介绍。这样用户在搜索时只要指明类型,就可使搜索范围缩小很多,提高效率。不过,如果用户本身输入的类型不对,那么就可能遗漏某些重要的信息源。分类式搜索引擎中最具代表性的就是 Yahoo。

对于软件下载搜索引擎同样需要做到这点,能够显示出软件所在的服务器和路径。一旦用户觉得通过 WWW 下载速度缓慢时,可以通过文件传输服务 FTP 直接登录到该服务器进行文件下载。

4. 能区分搜索结果的相关性

搜索引擎对于分类结果的排序上应该有所讲究。由于屏幕有限,搜索到的结果不可能同时出现在用户面前,很多时候还需要用户不断地“翻页(Next ...)”,因此最好把搜索到的结果按照其和用户关键字的相关程度排列。相关程度最大值为 1,如果为 0,则意味着该文档中根本没有关键字出现。在搜索引擎尚未完全智能化的今天,比较可行的方法就是按照关键词出现的频度,来排列搜索结果。关键字出现越多,频度越高,文档和关键字的相关程度也就越大,这样的文档理应出现在越靠前的地方。但是,必须注意的是,相关性越高的文档并不一定就是用户最需要的文档。

5. 检索方法多样,检索手段完备

有些性能完善的搜索引擎不仅能检索 Internet 上的文献,而且还可以查找公司和个人的信息;不仅能检索 Web 页面,还能提供对新闻组内文章的查找;不仅能对输入的单词、词组和句子进行检索,还能指定多个单词之间的逻辑匹配及其位置关系;不仅能以词语查询有关主题的页面信息,也能以特定的域名、主机名、URL 网址等查找有关的信息;此外,还能对被检索的文献进行语种、日期、大小等方面的限制。

6. 迅速

搜索到一个网站的速度是衡量搜索引擎优劣的基本指标。对于一个搜索引擎而言,速度越快自然越得人心。没有人能离开搜索引擎,但是没有人上网是为了专门浏览搜索引擎

的。搜索引擎必须尽快把结果提交用户才能赢得信任。为了确保速度，不仅要有快速的内部数据库搜索软件，而且搜索引擎的页面不应该过于花哨或有太多的广告，以免在网络上载时影响速度。

7. 其他性能

搜索引擎其他的一些性能是：可维护、可更新性，稳定可靠、可容错，有迅速修复的机制、性能价格比高。

1.3 搜索引擎语法规则

搜索引擎的工作基于用户输入的关键字，而其检索的效率在很大程度上也和关键字有关。由于单个的关键字搜索往往不能得到好的搜索结果，因此通常搜索引擎都允许用户按照自己提供的一系列语法来组织自己的关键字，以使搜索结果最优。

之所以要在这里就把语法规则告诉读者而不是对每一种搜索引擎分别论述，是因为目前的搜索引擎在语法规则上有相当一部分是相同的。虽然不是所有的搜索引擎都支持下面提到的这些语法规则，但是出入通常不会很大。掌握了这些语法规则，可以说就掌握了搜索引擎的统一语言，在任何搜索引擎上都不至于手足无措。另外，由于搜索引擎自身也在不断完善，因而不可能尽述一个搜索引擎的语法规则，也许还会有新的语法规则出现。

1. 使用逻辑操作符

搜索引擎中的逻辑操作符有：AND(与)，OR(或)，NOT(非)。

- (1) AND 在逻辑上表示“与”。如果用它来连接两个词汇，表明要求两个词汇都必须包括在搜索结果中。AND 可以简写作“&”。

【例】计算机 AND IBM

这将找到所有既包含“计算机”又包含“IBM”的相关文档。

- (2) OR 在逻辑上表示“或”。如果用它来连接两个词汇，表明或者这个词汇或者那个词汇要出现在搜索结果中，当然也可能两个词汇都出现。OR 可简写为“|”。

【例】计算机 OR IBM

其查询出来的结果显然要比 AND 查询出来的多。因为包含“计算机”但不包含“IBM”的文档和包含“IBM”但不包含“计算机”的文档都将被统计作为结果。

- (3) NOT 在逻辑上表示“非”，如果用它来连接两个词汇，那么表示搜索含有 NOT 之前的词汇的文档，但是排除里面含有 NOT 之后那个词汇的文档。NOT 可简写作“!”。

【例】计算机 NOT IBM

这时，只含有“计算机”而不含有关键字“IBM”的文档将作为结果传送回来。

在逻辑中以上 3 个符号是可以相互嵌套的，即可以采取一定的顺序将它们组合起来使用。比如要搜索计算机行业的营销状况，但是对 IBM 的情况不关心，可以这样写搜索条件：

计算机 AND 营销 NOT IBM

或者, 计算机&营销! IBM

 **注意:** 逻辑符号在解释时有优先级。通常 AND 和 NOT 要在 OR 之前进行解释。另外, 建议用户使用英文符号而不要使用简写, 因为支持后者的搜索引擎实在太少。

2. 使用“+”、“-”和通配符“*”

- (1) “+”、“-”都可以看成是一种强制性的符号, 它们可以用来强调某个词汇必须出现在搜索结果中, 或者某个词汇必须不出现在搜索结果中。它们应当写在要强调的词汇之前, 并且和词汇之间不能有空格。

 **注意:** 使用“+”来要求关键字出现在搜索结果中, 并不是说一定要出现在超链接的地址中。很可能它是出现在描述该地址的索引文档中。

【例】如果要查找联想公司的计算机产品但是不希望找“天琴”系列, 可以这样写:

+联想 -天琴

“+”、“-”号也可以配合上面讲过的逻辑符号使用。

【例】如果希望查找 IBM 的芯片但不要 Pentium MMX200 型号的, 可以这样写: 芯片 AND IBM -MMX200

- (2) “*”通配符的使用方法同在 DOS 操作系统下使用该通配符类似, 它可以用来指代词汇的某一部分。例如 com*, 可以代表 computer, communication, company, ……。这种语法通常只适用于英文文档搜索, 而且目前中文搜索引擎似乎还没有支持这种搜索方式的。

3. 使用NEAR操作符

有些搜索引擎提供 NEAR 操作符。其控制语句为 NEAR/n。n 是大于 1 的整数。它用于寻找在一定区域范围 n 内出现关键字的文档。这种操作符的使用显然可以很好地进一步完善“相关性”。它从统计关键字出现的频度到检查关键字相隔的距离, 用户可以输入一个范围, 比如 n 取值 100, 则检索关键字之间的间距最大不超过 100 个单词。如果关键字之间的距离越小, 那么其在结果中的排位就越靠前。

【例】输入: computer NEAR/100 game

则意味着查找 computer 和 game 之间间隔不大于 100 个单词的文档。

4. 使用逗号、括号和引号

- (1) 逗号的作用可以看成和“OR”的作用一样。对于用逗号分隔开的几个关键字, 在搜索时相当于是按照“或”的规则, 至少保证一个出现在结果中。不过, 它遵循的是“越多越好”的原则, 查询到的关键字越多, 则排位越靠前。

【例】编辑, 出版, 发行

查询时同时包含“编辑”、“出版”和“发行”的文档将出现在前面。

- (2) 括号的作用是为了在多种符号组合使用时调整优先级。用括号括起来的部分将具有最高优先级。下面将给出具体的例子:

【例】计算机 OR 电脑 AND 杂志 OR 游戏

提交后将搜索出或者含有“计算机”的文档,或者含有“电脑”和“杂志”的文档,或者含有“游戏”的文档,或者是包含以上两项或三项关键字的文档。比如关于“儿童游戏”的文档就可能出现在结果中。

【例】(计算机 OR 电脑)AND(杂志 OR 游戏)

由于括号调整了逻辑运算顺序,用户在提交后,实际查询的关键字就是“计算机杂志”,或者“计算机 游戏”或者“电脑 杂志”,或者“电脑 游戏”。

- (3) 引号的使用目的,是为了告诉搜索引擎将几个关键字作为一个完整的组合字符串进行搜索。比如要查找关于电子杂志方面的信息,就可以输入“electronic magazine”(带引号)。这样搜索引擎就把 electronic magazine 作为一个整体来认读。如果不加引号,根据通常的默认规则,空格作为 OR 解释,那么搜索引擎就会分别找到含有 electronic 和 magazine 这两个关键字的网页,与用户要求相去甚远。



注意: 有些搜索引擎具有良好的认读性能。比如 Yahoo, 即使用户输入的是不带引号的 electronic magazine, 引擎搜索时可能把两者分别检索, 但是在结果排列上, 会把两个词位置最靠近的文档放在最前面。这就保证了因为用户忽略, 而输入关键字又确实具有“整体认读”性时出现偏离主题的情况。但北大的“天网”就不具这样的性能, 它会把关键字按照 OR 的关系进行搜索处理。

5. 使用空格

当输入汉字作关键字的时候, 不要追加不必要的空格。因为在搜索引擎中, 空格被认为是一种特定的操作符, 其意义或许解释为 AND, 或许解释为 OR, 这由不同的引擎决定。

6. 使用连字符“-”

连字符在英语当中比较常用, 比如 Net-revolution(互联网革命), e-business(电子商务)。但是如果在 Yahoo 中输入 e-business, 恐怕搜索到的只是带 e 或者带 business 的网页。如果要纠正, 那么就应该使用“e-business”。

1.4 搜索策略与技巧

用户都希望自己的搜索有效、快捷、准确, 但是往往搜索到的内容不尽人意。这固然和搜索引擎本身的智能有关, 但也和用户自己搜索时的策略有关。以下一些提示可以帮助用户更有效地搜索需要的信息。

- (1) 不同目的的查询应当选用不同的搜索引擎。虽然搜索引擎有许多, 但是不可能每个搜索引擎都是绝对完善的, 而可能它的弊端恰好为另外的搜索引擎所弥补。有些搜索引擎还具有很强的专业功用。用于软件下载服务的 FTP 搜索引擎就和一

般的 http 文献搜索引擎不同。

- (2) 适当地运用语法规则进行组合搜索。搜索引擎的统计数据表明,很多用户只输入一个词来进行查询,这会带来很多不必要的匹配。为此,适当的使用前面所提到的语法规则进行组合搜索,可以帮助用户得到更为精细准确的结果。只要关键字挑选得越精确恰当,搜索结构就越好。但是没有必要使用太多的逻辑组合,使用搜索引擎毕竟不是做逻辑题,太多的组合可能使用户难以明确关键字之间的层次。可以利用搜索引擎提供的“层次搜索”来获得在现有结果上的进一步搜索,实现用户的要求。
- (3) 如果不能准确地确定要搜索什么,而只是为了搜索一类主题,那么最好使用 Yahoo 一类的目录式分类搜索引擎。这时用户不是直接输入关键词,而是要首先进入各种分类目录,然后再在适当的位置进行分类目录内的查询。
- (4) 尽可能地缩小搜索范围。许多站点允许仅在 Web、新闻组或者 E-mail 地址中搜索,或者就是允许进行地域性搜索。
- (5) 灵活运用短语。由于电脑的智能有限,如果确实希望电脑把关键字当作一个整体来搜索,最好使用引号来告诉搜索引擎,要搜索的关键字应该被看成是一个短语。
- (6) 大小写的敏感性。所谓的敏感性是指搜索引擎对于大写或者小写单词能不能区分寻找并搜索到相应的小写或者大写单词。若输入的都是小写字母的单词,则对大小写不敏感,这时很可能丢失含大写单词的超链接;而输入大写字母的单词,则对大小写都敏感。自然,对于中文搜索引擎来说,不存在这个问题。
- (7) 使用交叉话题,而尽量减少包容话题。比如为了知道“克隆(clone)”技术对社会以及政治产生的影响,如果仅仅输入“clone”,搜索引擎可能会找到一大堆介绍 clone 的学术性文章。如果改以查询“clone AND politics”或者“clone AND society”,这样找到的就可能会是所需要的内容了。另外,搜索范围如果有包容关系,那么在搜索时实际是范围小的那个不起作用。比如“science fiction OR fiction”,这时前一项实际是不起作用的。因此用户输入的每个词最好都有搜索的价值。
- (8) 掌握搜索引擎的特性。不同的搜索引擎在语法规则和搜索要求上会有些差异。因此为了使用好一个搜索引擎而阅读它的帮助是比较明智的工作。通常搜索引擎会含有 Options、Advanced search 和 Help 几个选项,用户可以在搜索的空闲看看它们,以便熟悉所使用的搜索引擎的特性。

1.5 中文搜索引擎介绍

下面开始介绍 WWW 上常用的中文搜索引擎。搜索引擎的基本知识在前面都已做了阐述,因此这里不打算浪费读者的时间。对于各类中文搜索引擎,将突出陈述其特色,而不深入探讨其一般的特性。