

实用

网络搜索

邢志宇 周明杰 武二伟 著

zhongsou.com
中国搜索

西安地图出版社

天网·神州

责任编辑:王兴华

封面设计:艺林

ISBN 7-80670-886-3



9 787806 708866 >

ISBN 7-80670-886-3/G · 318

定价: 36.00元

实用网络搜索

邢志宇 周明杰 武二伟 著

西安地图出版社

图书在版编目 (C I P) 数据

实用网络搜索 / 邢志宇, 周明杰, 武二伟著. — 西安:
西安地图出版社, 2005

ISBN 7-80670-886-3

I. 实... II. ①邢... ②周... ③武... III. 计算机网络
—情报检索 IV. G354.4

中国版本图书馆 CIP 数据核字 (2005) 第 134004 号

实用网络搜索

邢志宇 周明杰 武二伟 著

西安地图出版社出版发行

(西安市友谊东路 334 号 邮政编码 710054)

新华书店经销 许昌市裕达印刷有限公司印刷

850 毫米×1168 毫米·1/32 13.25 印张 355 千字

2005 年 12 月第 1 版 2005 年 12 月第 1 次印刷

印数: 0001—2000

ISBN 7-80670-886-3/G·318

定价: 36.00 元



邢志宇，1959年9月出生，山东掖县人，中国民主同盟盟员，中国图书馆学会会员，河南省许昌市图书馆副研究馆员。1982年毕业于许昌师范专科学校英语专业，1992年毕业于武汉大学图书情报学院图书馆学专业（本科函授），在省级以上学术刊物（会议）发表（宣读）论文40多篇，长期从事业务辅导、参考咨询和网络导航工作。



周明杰, 1973年4月出生, 河南省平顶山市人, 平顶山市图书馆馆员, 河南省财经学院会计系毕业, 本科学历, 学士学位, 河南省图书馆学会会员。发表图书馆学术论文11篇, 主持参与市、厅级社科重点调研课题6项, 获市优秀社科调研课题一等奖2项、二等奖2项。



武二伟 (1972), 河南省许昌县人, 中共党员, 1996年6月毕业于郑州大学信息管理系图书馆学专业。1992年至今在平顶山工学院图书馆工作, 现任平顶山工学院图书馆期刊部副主任、馆员。多年从事信息检索、信息组织与加工等方面的研究。独立撰写并正式发表学术论文8篇, 主、参编专业著作2部, 主持并参与市、厅级以上科研课题7项, 获市、厅级以上科技成果奖6项。

内 容 提 要

本书从实用的角度出发,阐述了网络搜索的基本理论和一般方法,分别详细、概要地介绍了 10 种常用搜索引擎和 300 多种元搜索引擎与专用搜索引擎,例释了典型的搜索技巧,详述了主要和有代表性的学科与主题网络资源。

本书为作者长期从事虚拟参考咨询服务(VRS)和网络导航工作的经验总结和成果汇集,具有较强的实用性,既可作为学习网络搜索的实用教程,也可作为参考工具书使用。适用于图书馆参考咨询人员、学术研究人员、大中专师生和广大的网络搜索爱好者。

网络搜索与搜索引擎研究

(代序言)

顾名思义,网络搜索就是在互联网上进行的信息搜寻。广义的网络搜索可以包括网站浏览式查找,但作为信息挖掘、图书馆学和情报学研究内容的网络搜索是狭义和主流的“网络搜索”,即利用搜索引擎进行的网络信息的搜寻与查找。

我国上网用户人数达到 10300 万,搜索引擎已成为仅次于电子邮件的第二大网络应用。凡有过网络经历的人,无不对 Google、Yahoo、百度、搜狗、中搜、一搜、北大天网等耳熟能详,无不有过网站、网页、新闻、图像、游戏、MP3、FTP 等搜索体验。网络的发展造就了搜索引擎,搜索引擎使人们能够快捷和有效地利用网络。无论对一般网民和工程技术与学术研究人员来说,搜索引擎已成为不可或缺的网络信息检索工具。

对商家来说,搜索引擎是一种赢利的产品,而对使用者而言,搜索引擎则仅仅是一种检索工具。作为产品,搜索引擎公司和搜索技术提供商要研制和创新搜索技术;作为工具,使用者要了解搜索引擎的功用、性能,掌握其使用方法和技巧。利用搜索引擎的目的不同,构成了搜索引擎研究的不同群体和对搜索引擎不同角度不同侧重的研究。

搜索技术研究的主体是搜索引擎的创制者和搜索技术提供商,它们多为专业公司 and 研究团体,拥有专职的研发队伍,专门从事网络信息挖掘理论研究、搜索引擎的开发、研制和搜索技术的改进、创新,靠建立搜索网站或提供搜索技术生存和赢利。如 Google 拥有独特的

链接评价体系 and PageRank 专利技术; Yahoo! 在收购了 Inktomi、Overture、Fast、AltaVista、Kelkoo、3721 等公司后,2004 年 2 月断然与 Google 决裂,推出融合各旗下公司搜索技术精华的自有搜索技术(YST),并开始成为搜索技术提供商;Metacrawler 开元搜索引擎先河,它的价值不仅在于同时调用多个独立的搜索引擎分别进行搜索的创意,更体现在处理各搜索引擎的查询结果,提交统一格式结果页面的独特技术;AltaVista 第一个开发了支持自然语言的搜索技术;Ask Jeeves 开创智能答询搜索引擎模式;About.com 推出垂直主题搜索引擎模式;hobomo 独创“MFGS 闪电搜索”技术和“变异再生呈现”技术;百度拥有“超链分析”专利技术;中搜开发的“划词搜索”技术;一搜依托 YST 的本地化搜索;HillTop 算法专利技术;vivisimo 公司的搜索结果自动个性化聚类技术;Blinkx 公司推出号称“可以搜索 200 余种媒介类型,包括了 Word,Excel,PowerPoint 和 Acrobat PDF 等格式的文档,以及各种音频与视频格式”的全新的“占卜式”搜索技术;搜狐 2004 年 8 月 3 日推出了自主研发的第三代搜索引擎——“搜狗”……由于搜索引擎具有较大的赢利空间和诱人的前景,搜索技术无时不在研发和创新之中,搜索引擎商无不具有独门产品和自主知识产权的特色技术。

网络发展的需要和自身生存的压力,促使搜索引擎商不断对搜索引擎进行改进和升级,推出新的功能,以满足不断增长的搜索需求,争取更多的用户。以 Google 为例,自 1998 年 9 月 7 日问世以来,它从未停止过其搜索技术的创新:1999 年 2 月增加短语搜索功能,2000 年 7 月 27 日推出高级搜索功能,8 月 16 日增加布尔逻辑符“OR”检索,2001 年 1 月 31 日增加 PDF 搜索,10 月 15 日增加分类目录搜索,2002 年 10 月 12 日增加购物搜索,2003 年 8 月 13 日 Toolbar 2.0 新版问世,8 月 20 日增加单词释义功能,11 月 6 日推出 Deskbar,2004 年 1 月 13 日推出区域编码、产品代码、航班信息、车辆确认号码以及美国邮政监管号码等搜索,3 月推出本地搜索,4 月

推出拼音汉字自动转换功能,6月24日推出中英文字典、天气、股票、邮编区号、手机号码等实用搜索功能,7月推出“实名通”,9月9日推出新闻简体中文版,10月6日增加图书搜索,11月17日推出学术搜索,2005年2月23日推出电影搜索,3月7日推出桌面搜索中文版,4月20日推出 My Search History 个性化功能,8月24日推出 GoogleTalk 即时信息和网络语音电话服务,9月推出中文本地搜索和 Blog 搜索……Google 几乎每月都有新的变化,每年都有重大的改进和创新,以功能的不断改进、增强和增加,满足用户不断提高和变化的需求。其他搜索引擎无一不是孜孜以求,不断地创新技术,增强功能。为适应个性化的搜索需求和发展趋势,各搜索引擎争相展现技术优势,纷纷抢滩桌面和工具栏,一时间几乎所有的搜索引擎都推出了称谓不同的 toolbar 或 deskbar。然而,尽管搜索技术研究竞争激烈,新的搜索技术和搜索功能不断问世,但真正“智能化”和“个性化”的搜索技术还仍在人们的期待之中。

二

搜索方法研究属于应用研究,对搜索引擎的推广使用有着直接的影响,与搜索引擎的普及和利用水平相辅相成。网络检索的需要催生了搜索引擎,搜索引擎使用方法的研究极大地促进了网络检索技能和网络资源利用水平的提高。

搜索方法研究的群体最为庞大,广大的网络搜索爱好者是其中的生力军,高校图书情报院系和图书情报部门的科研人员是网络搜索方法研究的主力军和骨干,网络搜索教学网站和搜索论坛是普及搜索知识、学习搜索方法、交流搜索经验和探讨搜索问题的主要平台,搜索引擎普及与研究网站和图书情报专业期刊是发表研究成果的主要载体。

较有影响的网络搜索教学网站有:extreme searcher (<http://extremesearcher.com/handbooklinks.html>)、Finding Information on the Internet: A Tutorial (<http://www.lib.berkeley.edu/TeachingLib/>)

Guides/Internet/FindInfo.html)、Anne Clyde Weblog (<http://www.hi.is/~anne/>)等。研究搜索方法较为知名的论坛有: Search Engine Watch 的 forum (<http://forums.searchenginewatch.com/>)、Search Guild 的 forum (<http://www.searchguild.com/>)、中文搜索引擎指南的“搜索论坛”(<http://www.sowang.com/bbs>)、中国生命科学论坛的“文献与检索技巧”(<http://www.bioon.net/list.asp?boardid=17>)、丁香园医学论坛的“文献检索知识交流版”(<http://www.dxy.cn/bbs/post/page?bid=6&sty=1&age=0>)、搜狗说吧——搜索引擎吧、百度贴吧——搜索引擎吧等等。此外,搜索引擎普及与推广网站一般也都设置专栏,登载搜索引擎功能与特点、各种网络资源搜索方法与技巧等方面的文章,其中影响较大的有中文搜索引擎指南网(<http://www.sowang.com/>)、《搜索研究院》(<http://www.9238.net/>)、搜索引擎直通车(<http://www.se-express.com/>)等等,这些都是不可多得的研究搜索引擎、搜索技巧和学习专业文献检索的门径。

近年来,高校《文献检索》课程均增加了搜索引擎和网络信息检索的相关内容,对普及网络搜索知识起到了积极的推动作用。图书情报界同仁对网络搜索的理论和实践从不同的角度进行深入的研究和系统的总结,各具特色的研究成果亦多有问世。在维普《中文科技期刊数据库》中进行关键词检索,“搜索引擎”检索出文章 3249 篇,“网络信息检索”407 篇,“internet+ 文献检索”54 篇,以“搜索引擎”、“网络信息检索”为关键词,从国家图书馆《中文普通图书数据库》分别检索出有关专著 23 种和 14 种。综观现有研究成果,搜索方法研究的范围比搜索技术研究要宽泛的多,涉及搜索引擎的检索功能、检索语法、检索策略、检索词语、检索技巧、检索工具的评价与选择、各个学科与各种类型网络资源检索等诸多方面。我们有理由相信,随着社会信息意识的增强,网民人口将快速增长,搜索引擎的使用率将逐步提高,搜索方法的研究将会受到人们更多的关注。

网络搜索离不开搜索引擎,对搜索引擎的了解和掌握影响着网络搜索的质量和效率。了解搜索引擎,很重要的一个方面就是及时获得搜索引擎的诸如版本升级、性能优化、算法改进、功能增加等动态信息。搜索引擎的动态信息有官方发布和研究者或使用者发布两种方式。前者多以新闻形式出现,可以通过新闻搜索及时获得,或在有关网站进行专题新闻查阅。设立搜索引擎专题新闻栏目(如: Search Engine News、Search Engine Update、搜索快报、搜索引擎业界动态等)的主要网站有: Search Engine News (<http://www.searchengineguide.com/searchengineneews.html>)、Search Engine Showdown (<http://www.searchengineshowdown.com/>)、ResearchBuzz (<http://www.researchbuzz.com>)、Search Engine Watch (<http://searchenginewatch.com/>)、Pandia Search Engine News (<http://www.pandia.com/sew/index.php>)、中文搜索引擎指南网 (<http://www.sowang.com/>)、搜索引擎直通车 (<http://www.se-express.com/>)等,从中可以进行专题新闻的系统浏览和回溯查找;后者多以 Blog 为载体,信息源较为广泛,有的是搜索引擎的研制者专注于某一搜索引擎的 Blog,更多的是搜索专家或搜索爱好者的个人 Blog。受人关注的此类 Blog 有: Google Blog (<http://www.google.com/googleblog/atom.xml>)、Yahoo! Search blogh (<http://www.ysearchblog.com/index.xml>)、msnsearch's WebLog (<http://blogs.msdn.com/msnsearch/rss.aspx>)、Ask Jeeves Blog (<http://blog.ask.com/index.rdf>)、SEW Blog (<http://blog.searchenginewatch.com/blog/>)、搜索引擎研究 (<http://www.wespoke.com/>)、Google 研究院 Blog (<http://www.ooone.com/blog/>)等等。通常情况下,搜索引擎不会发布其测试功能和日常改进的官方信息,而有关专家的个人 Blog 则会对搜索引擎的细微变化有敏感的反映,利用 Blog 搜索引擎和 RSS 订阅可以及时获得搜索引擎的非官方动态信息。

我们已跨进信息社会的门槛,互联网已成为人类共享的信息平台,并正在影响和改变着我们的生活。然而,网海浩瀚资源丰富信息

茫茫,从中获取我们所需要的特定的资源或信息,没有专门的检索工具是不可想像的。网络搜索与搜索引擎研究,对提高社会的网络应用水平,培养和强化信息意识,进而促进知识生产和知识创新,有着重要的意义和影响。“工欲善其事,必先利其器”,了解搜索引擎的功能、特点,探讨和研究各种网络资源的检索策略、方法和技巧,熟练掌握和使用搜索引擎,已成为现代人必不可少的基本技能。

邢志宇

2005年9月18日

目 录

第一章 搜索入门	(1)
第一节 明确搜索意图	(1)
第二节 制定搜索策略	(5)
第三节 选择搜索工具	(8)
第四节 构建检索式	(11)
第五节 搜索结果评价	(17)
第二章 搜索引擎	(20)
第一节 常用搜索引擎	(21)
一、网络搜索巨鳄--Google	(21)
二、分类搜索引擎鼻祖--Yahoo!	(39)
三、著名的多媒体搜索引擎--AltaVista	(45)
四、最受欢迎的元搜索引擎--Dogpile	(50)
五、最大的中文搜索引擎--百度	(55)
六、第三代互动式搜索引擎--搜狗	(63)
七、全球领先的中文搜索引擎--中国搜索	(71)
八、雅虎的中国搜索门户--一搜	(77)
九、中国的 ODP--网易	(81)
十、最大的 FTP 搜索引擎--天网千帆	(84)
第二节 元搜索引擎	(86)
一、中文元搜索引擎	(88)

二、英文元搜索引擎.....	(89)
第三节 专用搜索引擎.....	(93)
一、多媒体搜索引擎.....	(93)
二、新闻搜索引擎.....	(97)
三、“看不见的网页”与数据库搜索引擎.....	(100)
四、FTP 搜索引擎.....	(103)
五、购物搜索引擎.....	(109)
六、少儿搜索引擎.....	(113)
七、MP3 搜索引擎.....	(117)
八、歌词搜索引擎.....	(122)
九、域名搜索引擎.....	(125)
十、软件搜索引擎.....	(127)
十一、论坛搜索引擎.....	(130)
十二、博客搜索引擎.....	(133)
十三、游戏搜索引擎.....	(135)
十四、文学作品搜索引擎.....	(139)
十五、答案搜索引擎和咨询服务系统.....	(140)
第三章 搜索技巧.....	(145)
一、查找中文网络报刊资料.....	(145)
二、企业情报调研中的搜索技巧.....	(149)
三、图像搜索的途径与方法.....	(155)
四、查找某一国家(地区)组织、机构的电邮与电话号码.....	(160)
五、查找地方实用信息.....	(162)
六、查找“海外客户资料”.....	(163)
七、查找商品信息.....	(165)
八、找网络书签或个人收藏夹.....	(166)

九、查找歌曲·····	(167)
十、查找歌词·····	(168)
十一、查找个人信息·····	(170)
十二、查找“明星”资料·····	(171)
十三、查找下载地址·····	(172)
十四、查找论文·····	(173)
十五、查找范文·····	(174)
十六、查找免费电子书·····	(175)
十七、查找解决问题的方法·····	(177)
十八、查找网站 Logo·····	(177)
十九、查找英文缩写·····	(178)
二十、查找注册码·····	(180)
二十一、查找论坛中的帖子·····	(181)
二十二、查找 Blog 中的文章·····	(183)
二十三、查找英文软件的中文使用说明·····	(184)
二十四、查找免费电影·····	(186)
二十五、查找网站地址和联系方式·····	(188)
二十六、找语词的最早出处·····	(190)
二十七、查找同类资源·····	(191)
二十八、查找数据库密码·····	(191)
二十九、查找讲座与讲义·····	(192)
三十、查找免费代理服务器列表·····	(193)
三十一、查找网站的 flash 首页和动态网页·····	(194)
三十二、查找国外网站的中文网页和国内网站的英文网页·····	(195)
三十三、查找中、英文课件·····	(197)

三十四、查找英汉互译·····	(199)
三十五、查找电话号码归属方位·····	(200)
三十六、查找成语与成语反查·····	(201)
三十七、查找试题与答案·····	(202)
三十八、查找谜底、对联·····	(203)
三十九、查找烹饪方法·····	(204)
四十、查找药品功效与使用方法·····	(205)
第四章 网络资源·····	(206)
第一节 网络资源调查·····	(206)
一、关键词调查·····	(208)
二、分类调查·····	(209)
三、链接调查·····	(215)
第二节 常用学科资源·····	(217)
一、哲学·····	(218)
二、宗教·····	(222)
三、社会科学(综合)·····	(224)
四、政治学·····	(226)
五、社会学与社会工作·····	(229)
六、经济学·····	(232)
七、教育·····	(235)
八、图书馆学、情报学·····	(238)
九、语言学·····	(241)
十、文学·····	(244)
十一、艺术(综合)·····	(247)
十二、美术·····	(248)
十三、摄影·····	(250)