

# 信息搜索技巧 Search Engine



南开大学出版社



*INTERNET* 随身系列

# 信息搜索技巧

翁智泓 原著  
佚民 改编

南开大学出版社

## 信息搜索技巧

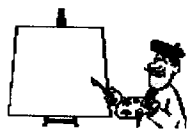
---

|             |                                                     |
|-------------|-----------------------------------------------------|
| <b>原著</b>   | 翁智泓                                                 |
| <b>改编</b>   | 佚民                                                  |
| <b>出版者</b>  | 南开大学出版社<br>地址:天津市八里台南开大学校内<br>邮编:300071 电话:23508542 |
| <b>法律顾问</b> | 庞标                                                  |
| <b>发行所</b>  | 新华书店天津发行所发行                                         |
| <b>承印商</b>  | 南开大学印刷厂印刷                                           |
| <b>版次</b>   | 1998年1月第1版                                          |
| <b>印次</b>   | 1998年1月第1次印刷                                        |
| <b>开本</b>   | 787×1092 1/32                                       |
| <b>印张</b>   | 6.75                                                |
| <b>字数</b>   | 144千                                                |
| <b>印数</b>   | 1-5000                                              |
| <b>ISBN</b> | 7-310-01153-8/TP·72                                 |
| <b>定价</b>   | 13.00元                                              |

---

本书由台湾立威出版股份有限公司授权南开大学出版社  
在大陆地区独家出版发行  
书中述及各商标均属原注册公司所有

# 序



每次连上网络，心里总有份悸动，原因无它，只是赞叹现今信息的进步，没想到从这些冰冷的机器中竟可连到世界上的任一角落……。

当我们在网络上浏览时，可曾迷失在浩瀚无边的信息中？除了突破空间的限制之外，网络也常常让人忘了时间，而要如何让自己除了满足玩心之外更可以得到一些真正的东西，这些都是我们要好好思考的。

如何在 Internet 上高效地搜索到您所需要的信息呢？本书针对一些常见的搜集站

点的实际应用作了深入浅出的说明。这些站点的搜索引擎各有特色，可说是重点所在。尽管目前多数搜索引擎的界面未能直观地显现其应用要点，但只要花半小时熟悉一下其功能选项和关键词逻辑运算等技巧，即使一般用户也会立刻成为事半功倍的搜索高手。

信息诚如海洋般浩瀚，而 Web 从来就是漫无头绪。不过读完本书之后，您将发现人们总是努力将一切回复有序。搜索引擎使信息唾手可得，面对 Internet，您会凭添几分指点江山的豪情。

# 目 录

## 第一章 WWW 上的检索服务器

- 1-1 搜索引擎的起源 ..... 3
- 1-2 第一个搜索引擎 Archie ..... 6
- 1-3 如何用 Web 浏览器连上搜索引擎 ..... 7

## 第二章 Yahoo!搜索引擎

- 2-1 Yahoo!的起源 ..... 12
- 2-2 如何快速进入 Yahoo! ..... 13
- 2-3 功能说明 ..... 14
- 2-4 其他功能 ..... 21
- 2-5 找不到的情况 ..... 24

### 第三章 AltaVista 搜索引擎

|                         |    |
|-------------------------|----|
| 3-1 AltaVista 的起源 ..... | 30 |
| 3-2 AltaVista 的特点 ..... | 31 |
| 3-3 搜索的设置 .....         | 33 |
| 3-4 搜索的显示与范围 .....      | 39 |
| 3-5 限制性的搜索 .....        | 42 |
| 3-6 复杂的搜索 .....         | 54 |

### 第四章 蕃薯藤 YamWeb 搜索引擎

|                               |    |
|-------------------------------|----|
| 4-1 蕃薯藤站点的起源 .....            | 71 |
| 4-2 如何快速连上蕃薯藤 .....           | 73 |
| 4-3 蕃薯藤的特点 .....              | 75 |
| 4-4 蕃薯藤的功能说明 .....            | 84 |
| 4-5 蕃薯藤的 News/BBS 精华区索引 ..... | 97 |

### 第五章 DejaNews 网络讨论组的搜索引擎

|                              |     |
|------------------------------|-----|
| 5-1 DejaNews 的起源 .....       | 104 |
| 5-2 如何快速进入 DejaNews .....    | 105 |
| 5-3 DejaNews 的功能 .....       | 106 |
| 5-4 强力搜索(Power Search) ..... | 133 |

## 第六章 其他搜索站点

|                         |     |
|-------------------------|-----|
| 6-1 Infoseek 搜索站点 ..... | 149 |
| 6-2 Excite 搜索引擎 .....   | 168 |
| 6-3 搜索精确度的再探讨 .....     | 186 |

## 第七章 搜索引擎的比较和展望

|                                  |     |
|----------------------------------|-----|
| 7-1 让 Surfing 不再只是 Surfing ..... | 193 |
| 7-2 搜索方法的评估 .....                | 196 |
| 7-3 搜索站点的比较 .....                | 199 |
| 7-4 搜索站点未来发展 .....               | 204 |

## 附 录

|                 |     |
|-----------------|-----|
| 各搜索引擎站点列表 ..... | 205 |
|-----------------|-----|





# WWW 上的 检索服务器



在 Internet 的流行方兴未艾之时，每个人都在关心着到底它将如何改变我们未来的生活。然而当我们一旦踏上 Internet 这列信息时代的快车，摆在我们面前的第一个问题就是如何在这样一个信息的海洋中找到自己所需要的信息和资料。在成千上万的站点中，已经有一些站点是为了帮我们快速找到所需资料而设立的，这些站点

就是我们所说的搜索引擎(Search Engine)。这些站点的设立确实带给我们很大的便利。但是这些站点究竟如何驾驭？每个人都可以利用它减少在网络上寻找信息所花费的时间吗？这就是我们要讨论的重点。

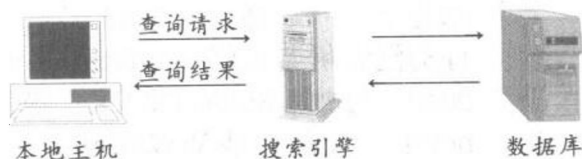
- 搜索引擎的起源。
- 以前搜索引擎是什么形式？
- 如何用 WWW 浏览器连上搜索引擎？

## 搜索引擎的起源

早在几年前, Internet 上的资源虽然五花八门, 但是我们都可以做到对于自己感兴趣的某些站点很了解, 而且因为常常上站, 尽管面对的是冰冷的文字模式, 但还是觉得很方便。

曾几何时, WWW 的诞生改变了 Internet 的命运, 文字模式变成了活泼生动的图形界面, 而且站点的增加速度更是惊人, 无法再以传统的记忆方法来应付与日俱增的站点数, 故搜索引擎应运而生。

搜索引擎是通过我们输入所需信息的关键词, 经由检索服务器处理内部数据库, 找到相关的资料并整理后再送出来, 通过 Internet 传到我们的本地主机。



搜索数据库内包括了已经整理好的由站点名、关键词等所组成的许多条记录，当然这些站点都自己设计了不同的算法来负责这些超大型的数据库的维护，使其能够全天候接待来自世界各地用户的访问，当然还包括如何去及时更新这些数据。

搜索引擎有两种，其功能和它当初所设计的用途有关，现就这两种分别说明如下：

- 一、第一种是直接分类，基于字母顺序、年代、地理或事件来分类，或是上述的合并，需要人们很辛苦地搜集资料、组织内容、编写庞杂的 HTML 文件。通常比较小的站点才采取这一种做法。
- 二、第二种是全自动搜集并做索引分类的方式，不需要花费大量人力投入，是未来发展的趋势。

搜索引擎的工作由两部分组成——收集和搜索。有一些很有趣的方法，如利用全自动漫游机器人(Automated Robot Wanders)、大蜘蛛(Spider)等工具，漫游于 Internet 中，大部分是 WWW 站点，也包括一

些 FTP 及 Gopher 站点，利用这些工具带回资源，然后将排序过的资料、索引来创建数据库。例如 Lycos 站点就是用的这种方法。

## 1-2

## 第一个搜索引擎 Archie

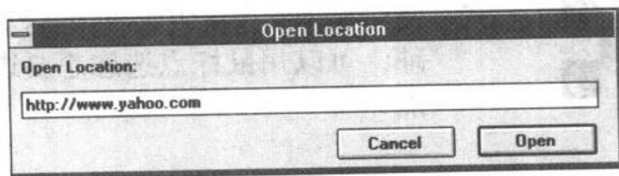
还记得 Internet 开始不久时我们所使用的第一个搜索的工具是什么吗？相信你我一定都用过，那就是 Archie。它是由 McGill 大学的一个小组开发的。我们以前很习惯地拿它来找我们所要的文件，而这些大部分都是 FTP 站点的一些文件，我们只要大概知道文件名。例如开头是 compu，它就可以帮我们找出文件名内含有 compu 的文件，并且把其所在的站点及目录都显示出来。

如同现在的搜索引擎，Archie 所搜索文件的方式当然也不是等到用户发出命令后才去把所有 FTP 服务器中的文件找一遍。因为这样做的代价是非常高的，而且也没有必要。它定期接受由其他站点送来的文件名，用加索引的方法来使查询者可以很快地找到所要的文件名。

## 如何用 Web 浏览器连上搜索引擎

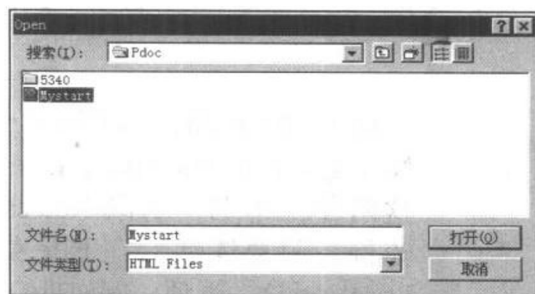
1-3

现在一般所用的 Web 浏览器大多是 Netscape Navigator 或 Microsoft Internet Explorer 这两种，本书不妨以 Netscape Navigator 为例。只要连线后在 File 菜单下选 Open Location 这一选项，例如我们要连至 Yahoo! 这个站点，在对话框中输入 `http://www.yahoo.com`，如图：



然后就可以进入 Yahoo! 这一个站点。要进入其他站点可以依此类推。本书后面的附录有其他站点的地址，可以按照上面的方法——把站点的地址输进去。或者你也可以把它们编辑成 HTML 文件存在本地硬盘上，进入 Web 浏览器后再把它调进来，方法如下：

- 一、进入Netscape, 在File菜单中选Open File 菜单项, 就会出现如下画面:



- 二、只要用鼠标选择该 HTML 文件并按下“打开”按钮, 这个文件就可以读进来。
- 三、此时画面会显示出一些搜索站点的地址, 可以用鼠标点选您想要进入的站点。





# 2

## Yahoo! 搜索引擎



Yahoo!站点的小档案：

- URL：<http://www.yahoo.com>
- 开发者：David Filo与杨致远(Jerry Yang)
- 收集的站点数：二十万个以上
- 搜索结果以三种方式显示：
  1. 在Yahoo!中对于该搜索字符串的分类。