



北京市高等教育精品教材立项项目

网络信息资源 管理与检索

饶伟红 主编

计算机网络技术
与应用系列教材

203

090



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

北京市高等教育精品教材立项项目

计算机网络技术与应用系列教材

网络信息资源管理与检索

饶伟红 主编

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书全面介绍了近年来网络信息资源管理与检索的最新理论、方法与技术,重点论述了网络信息资源的特点、类型、评价方法与标准,网络信息资源组织、管理的方式与方法,网络信息资源的有效检索与利用技巧,网络信息检索工具与系统的特点和使用方法,以及网络信息资源导引系统的建设与元数据管理技术等一系列网络信息资源管理与检索的核心问题。

本书内容新颖、重点突出、实用性强,面向网络信息资源用户与信息资源管理者,既可作为高等院校信息管理类及相关专业网络信息资源管理与检索相关课程的教材或教学参考书,亦可作为广大网络信息资源用户的实用参考书。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究。

图书在版编目(CIP)数据

网络信息资源管理与检索/饶伟红主编. —北京:电子工业出版社, 2004.7

(计算机网络技术与应用系列教材)

ISBN 7-121-00050-4

I. 网… II. 饶… III. ①计算机网络—信息管理 ②计算机网络—情报检索 IV. G252.7

中国版本图书馆CIP数据核字(2004)第061633号

责任编辑:李 岩

印 刷:北京牛山世兴印刷厂

出版发行:电子工业出版社

北京市海淀区万寿路173信箱 邮编 100036

经 销:各地新华书店

开 本:787×980 1/16 印张:13.5 字数:296千字

印 次:2004年7月第1次印刷

印 数:5 000册 定价:18.00元

凡购买电子工业出版社的图书,如有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系。联系电话:(010) 68279077。质量投诉请发邮件至 zltz@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

前 言

网络信息资源的快速发展,一方面极大地丰富和拓展了人类交流与获取信息的空间与渠道,另一方面也对网络信息资源的有效利用、网络信息资源的组织与管理提出了新的要求与挑战:网络信息资源的无限、无序、缺乏控制与用户所需的有序、有效、有针对性利用信息之间的矛盾,已成为当今信息资源组织、管理与利用领域的主要矛盾;网络信息资源区别于传统信息资源的动态性、多维性及数字化特性,使其与以往相比,在对信息资源的选择、评价、分类、标引、组织、管理及提供利用等方面存在很大差异,从而对传统的信息资源组织、管理与利用方式形成了强烈冲击。为此,确保网络信息资源的有效利用,根据网络信息资源自身的特点,运用现代信息技术、方法与手段,对网络信息资源实施加工、整理、控制和协调,使网络信息变无序为有序,解决网络信息无序扩张与有效利用之间的矛盾,为网络信息用户提供便利的信息获取途径与方法,既是广大网络信息资源用户极为关注和迫切需要解决的主要问题,也是我们编写此书的初衷。

本书是“北京市高等教育精品教材立项项目”之一——“计算机网络技术与应用系列教材”中的一本。本书面向网络信息资源用户与信息资源管理者,注重新颖性与实用性,重点介绍了网络信息资源的特点、类型、评价方法与标准;网络信息资源组织、管理的方式与方法;网络信息资源有效检索与利用的知识;网络信息检索工具与系统的特点与使用方法;网络信息资源导引系统的建设与元数据管理技术等一系列网络信息资源管理与检索的核心问题。在本书编写的过程中,注意将网络信息资源的组织、管理方法、信息检索方法与各种国际联机检索系统、各种网络信息资源检索工具的使用方法有机融为一体,使用户不仅能掌握网络信息组织、管理与获取的基本知识,同时能在较短的时间内,掌握管理与检索信息的方法与技能,提高用户在网络信息环境下的生存质量与发展水平。

全书共七章,其中第1章由符绍宏、饶伟红编写;第2章由饶伟红编写;第3、4两章由符绍宏编写;第5章由饶伟红、符绍宏编写;第6、7两章由饶伟红编写。全书由饶伟红总纂。

本书在出版过程中,得到出版社章海涛和李岩编辑鼎力相助,再此表示衷心感谢。

受作者认知水平所限,疏漏与错误之处在所难免,恳请同行专家与读者不吝赐教。

编 者

2004年4月

目 录

第 1 章 网络信息资源	1
1.1 互联网信息资源概述	2
1.2 互联网信息资源类型	3
1.3 互联网信息资源评价	7
第 2 章 网络信息资源的组织与管理	13
2.1 网络信息资源组织与管理的内涵及特征	14
2.2 网络信息资源组织与管理的发展	15
2.3 网络信息资源的组织方式	16
2.4 网络信息资源的组织方法——分类法与主题法	19
2.4.1 网络信息资源的分类组织法	19
2.4.2 网络信息资源的主题法系统组织法	22
2.5 网络信息资源的过滤	24
2.6 网络信息资源组织中的计算机系统体系结构	26
2.7 网络信息资源组织管理的理论基础与技术基础	27
2.7.1 网络信息资源组织管理的理论基础	27
2.7.2 网络信息资源组织管理的技术基础	29
第 3 章 网络信息资源的检索	33
3.1 网络信息检索概述	34
3.1.1 互联网信息检索的特点	34
3.1.2 互联网信息检索的一般方法	34
3.2 网络信息检索工具	36
3.3 网络信息检索技巧	41
第 4 章 互联网信息检索工具	45
4.1 目录型网络检索工具 (Web Directory)	46
4.1.1 概述	46
4.1.2 几种目录型检索工具	48
4.2 搜索引擎 (Search Engine)	58
4.2.1 概述	58
4.2.2 英文搜索引擎	61

4.2.3	中文搜索引擎	69
4.3	集合型网络检索工具——元搜索引擎 (Meta Search Engine)	77
4.3.1	概述	77
4.3.2	几种元搜索引擎介绍	79
4.4	专题型网络信息检索工具	92
4.4.1	社会科学类	92
4.4.2	教育类	94
4.4.3	医学类	97
4.4.4	工程类	101
4.5	其他类型网络信息检索工具	102
4.5.1	非 Web 资源检索工具	102
4.5.2	多媒体资源检索工具	110
4.6	网络信息检索工具性能指标与评价	117
第 5 章	因特网环境下的国际联机信息检索系统	121
5.1	联机信息检索系统	122
5.1.1	联机信息检索系统的组成	122
5.1.2	联机检索系统的工作原理	122
5.1.3	联机信息检索系统的检索步骤	123
5.1.4	联机信息检索系统的检索方法与检索式的构建	123
5.1.5	联机信息检索系统的服务方式	126
5.1.6	联机信息检索系统的起源与发展	126
5.2	国际联机检索系统与搜索引擎的比较	127
5.3	Internet 环境下的 Dialog 联机信息检索系统	128
5.3.1	Dialog 联机检索系统的检索渠道	129
5.3.2	Dialog 系统的 Web 检索平台	130
5.4	OCLC 的 FirstSearch 联机检索系统	144
5.4.1	FirstSearch 联机检索系统的特点	144
5.4.2	FirstSearch 联机检索渠道与应用环境	145
5.4.3	FirstSearch 联机检索系统的数据库	146
5.4.4	FirstSearch 的检索模式	148
5.4.5	FirstSearch 检索式的构建	150
5.4.6	FirstSearch 检索结果	151
5.4.7	FirstSearch 检索系统的一般检索步骤	151
5.5	Uncover 与 ingenta 联机信息检索系统	152
5.5.1	Uncover 联机检索系统的特点	152
5.5.2	Uncover 与 ingenta 联机检索系统的使用	152

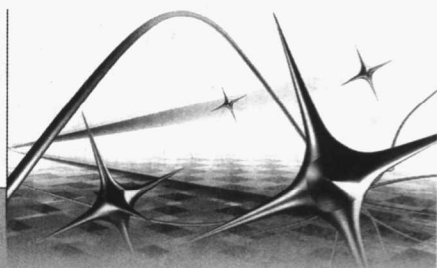
5.6	EI Village 与 EI Village 2	155
5.6.1	EI Village 的特点、结构与检索方式	155
5.6.2	Engineering Village 2 的特点与检索方法	158
5.7	SCI 联机信息检索系统	163
5.7.1	SCI 的特点	163
5.7.2	Web of Science 的检索	164
第 6 章	网络信息资源导引系统的建立与维护	171
6.1	导引系统建设目标	172
6.2	导引系统设计与实现	173
6.2.1	导引系统体系结构设计	173
6.2.2	导引系统的功能设计	173
6.2.3	导引系统网页设计	174
6.2.4	导引系统数据库设计	175
6.2.5	导引系统软硬件环境设计	177
6.3	导引系统的推广与宣传	179
6.3.1	搜索引擎注册	179
6.3.2	交互式链接	181
6.3.3	广告	181
6.3.4	电子邮件列表	181
6.4	系统维护与管理	182
6.4.1	系统的测试	182
6.4.2	系统的管理	183
6.4.3	系统维护	183
第 7 章	网络信息资源组织——Dublin Core 元数据	185
7.1	元数据概论	186
7.1.1	元数据的含义	186
7.1.2	元数据的类型	186
7.1.3	元数据的格式	188
7.2	Dublin Core 元数据	189
7.2.1	DC 的研究进程	189
7.2.2	DC 的特点	189
7.2.3	DC 的作用与功能	190
7.2.4	DC 的基本元素	190
7.3	元数据的创建与应用	191
7.3.1	元数据的创建	191
7.3.2	元数据的应用	193

附录 Dialog 检索系统数据库蓝页样例	196
思考与练习题	203
参考文献	205

互联网信息资源作为数字化、网络化信息的核心和集成，与传统的信息资源相比有很大的不同，了解其发展特点和独特的信息利用价值是有效利用网络信息的基础；互联网信息资源的种类有很多，除了占主要地位的 Web 信息外，还有 Telnet, FTP, Usenet Newsgroup, Mailing List 等；互联网信息资源的无限、无序、优劣混杂，使得发展网络信息资源的评价方法成为必要，掌握对互联网信息资源评价的相关标准和相关评价站点可以帮助用户有效地甄别、筛选出具有较高质量和利用价值的网络资源，提高利用互联网信息资源的效率。

第 1 章

网络信息资源



1.1 互联网信息资源概述

互联网是世界上最大的信息宝库,它已成为全球范围内传播和交流科研信息、教育信息、商业和社会信息的最主要的渠道。要想在这一浩瀚无边、变化多端而又鱼龙混杂的信息海洋中发现并查找出有利用价值的信息并不是一件易事。在以互联网为核心的电子信息环境下完成确实、有效的信息检索对于每一个检索者来说都是一项十足的挑战。而要进行有效的网络信息检索,首先就必须对互联网上信息资源的分布、种类和利用价值等有较全面的认识 and 了解。

互联网作为数字化、网络化信息的核心和集成,它与传统的信息媒体和信息交流渠道相比有很大的不同,其主要特点可概括为:

- 信息资源极为丰富,覆盖面广,涵盖了各学科领域,且种类繁多,几乎无所不包。
- 超文本、超媒体、集成式地提供信息,除文本信息外,还有图表、图形、图像、声音、动画等。
- 信息来源分散、无序,没有统一的管理机构,也没有统一的发布标准;且变化、更迭、新生及消亡等都时有发生,难以控制。

互联网提供了一种全新的交流信息和查找信息的渠道,它具有方便、及时、快速和交互性等特点,互联网信息资源在以下一些方面具有优越性和重要的信息利用价值:

- 价廉。它是一种比印刷品便宜的信息提供方式。不仅提供信息线索和著录信息,还提供有关信息的全文和原稿。
- 新颖、深入。互联网提供了获取非出版信息的丰富机会。如网上大量的灰色文献或边缘文献(Grey or Fringe Literature)——即在主流出版物渠道之外的文献,包括研究报告、调查采访、研讨会发言、笔记、项目计划报告和政策方针等。它们反映了许多研究成果背后的原始数据或第一手资料,大概是因为其内容太新或太专而未被纳入正式文献交流渠道。
- 广泛、直接交流。互联网扩大了人际交流的范围,提供了更多的直接交流机会。如参加用户网(Usenet)的新闻组(Newsgroup)、讨论组(Discussion Group)、邮件列表(Mailing List)等的讨论。还可在许多学者、研究人员、咨询专家的个人网页上发现其研究心得、教学演讲用的资料、演示、指南性的工具等,这是一种颇具个人特长的知识库,其参考价值应被重视。
- 非正式和自由发表。互联网提供了在正式的出版和发表渠道之外的、发表个人见解的空间,有较大的自由度,因而为新观点、不成熟的观点、未成定论的理论、假说、概念等提供了发表的园地。

总之,互联网信息资源不是传统信息资源的复制。互联网也不能取代传统的信息媒体和交流渠道,它是对传统信息资源和信息交流渠道的最令人振奋、最有力的补充。可以断言,对科研工作者来说,不掌握互联网这一信息工具,不利用互联网信息资源,就

很难达到相关领域的前沿, 很难取得具有国际先进水平的工作成果。

1.2 互联网信息资源类型

互联网信息资源包罗万象, 广泛分布在整个网络之中, 没有统一的组织管理机构, 也没有统一的目录。但按照其所采用的网络传输协议的不同, 可将互联网信息资源划分为以下几种主要类型:

1. 万维网 (World Wide Web, 简称 WWW 或 Web) 信息资源

WWW 信息资源指建立在超文本、超媒体技术的基础上, 集文本、图形、图像、声音为一体, 并以直观的图形用户界面 (GUI) 展现和提供信息的网络资源形式。WWW 是互联网的宠儿, 代表着互联网信息资源的主流。它使用简单、功能强大。自 20 世纪 90 年代问世以来, WWW 发展极为迅速, 它的超文本、超媒体特性使之在互联网信息资源提供、存储和检索等方面都独占鳌头。以下介绍几个与 WWW 相关的重要概念:

1) 超文本传输协议 HTTP

HTTP 是浏览器与 Web 服务器之间相互通信的协议, 即 WWW 客户机和服务器用于在网上传输、响应用户请求的协议。

2) 超文本标记语言 HTML

HTML 是一种专门编程语言, 规定和描述了文件显示的具体格式。也即 Web 服务器的信息是用 HTML 来描述的, HTML 文档由文本、格式代码和到其他文档的链接所组成。

3) 统一资源定位器 URL

URL 实际上是一个用以标识文档类型及其所在网络地址的字符串, 它的用途是用统一的方式指明互联网上信息资源的位置。URL 包括三部分: 所使用的传输协议, 服务器名称, 该服务器上定位文档的全路径名。URL 不仅用于 HTTP 协议, 还可用于 FTP, Gopher, Telnet 等协议。

4) 主页 (Home Page)

主页是在与任何一个 Web 服务器连接后所见到的第一页 (起始页)。它是该服务器入口处的 HTML 文件, 是进入该服务器和访问其所提供的各类信息资源的导引页。

5) 链接 (Link)

链接是 Web 页的元素, 它是指向其他信息资源的指针。

6) 浏览器

浏览器是一种应用于 WWW 的网络软件, 是一种驻留于客户端的应用程序。它不仅可用于与其他 WWW 服务器的连接, 其更主要的用途是帮助用户浏览、阅读和查找 WWW 信息资源。目前, 广为网络用户熟悉和使用的浏览器是 Netscape 和 Internet Explorer (IE)。它们不仅可以浏览文本信息, 还可显示与文本内容相配合的图片、图形和声音等。

WWW 浏览器的另一优势是可以使用多种协议进行工作。因此, 通过一个功能较强的浏览器就可访问 Gopher, Wais, FTP, News 等许多其他类型的网络资源。某些网络用户甚

至只使用 Netscape, IE 等浏览器就可满足大多数网络应用的需求。总之, WWW 信息资源是互联网信息资源的最主要、最常见的形式。

2. 远程登录 (Telnet) 信息资源

Telnet 信息资源指借助远程登录 (Remote Login), 在网络通信协议 Telnet (Telecommunication Network Protocol) 的支持下, 在远程计算机上登录, 使自己的计算机暂时成为远程计算机的终端, 进而可以实时访问、使用远程计算机中对外开放的资源。简而言之, 就是通过远程登录后, 可以访问、共享的远程系统中的资源。这些资源既包括硬件资源, 如超级计算机、精密绘图仪、高速打印机、高档多媒体输入/输出设备等; 也包括软件资源, 如大型的计算程序、图形处理程序、大型数据库等信息资源等。

访问 Telnet 信息资源的一般步骤是, 在 Telnet 命令中给出远程计算机的域名或 IP 地址, 在远程呼叫成功后, 键入自己的用户名和口令; 而某些提供开放式远程登录服务的计算机不需事先取得用户名及口令, 可以使用通用用户名 “public” 或 “guest” 进行公开登录; 在登录成功、建立连接后, 就可按给定的访问权限来访问相关资源——远程计算机上的软硬件、数据库等, 或使用该远程计算机提供的互联网其他信息服务, 如: E-mail, FTP, Archie, Gopher, WWW, WAIS, BBS 等。可见, Telnet 是一个强有力的资源共享工具。通过 Telnet 方式提供的信息资源主要是一些政府部门、研究机构对外开放的数据库; 主要商用联机检索系统, 如 Dialog, DataStar, Lexis-Nexis, OCLC, UMI 等也提供 Telnet 形式的连接方式, 进而检索其数据库, 不过需事先付费取得账号及口令。也有一些免费的系统, 如常见的由许多大、中型图书馆通过 Telnet 方式提供的联机图书馆公共检索目录 (OPAC, Online Public Access Catalog) 可以使用户在远程登录后, 联机检索该图书馆的馆藏目录, 了解其馆藏图书的书目信息。

3. FTP 信息资源

FTP (File Transfer Protocol) 是 TCP/IP 协议族中的一个, 是互联网使用的文件传输协议。其主要功能是在两台位于互联网上的计算机之间建立连接以传输文件, 完成从一个系统到另一个系统完整的文件复制。前面所述的 Telnet 可以使用户在远程登录后的联机状态下浏览、检索、利用远程计算机的资源, 但如想获取、拥有远程计算机中某些文件的复制, 则要利用文件传输服务 FTP。FTP 不仅允许从远程计算机上获取、下载文件 (Download), 也可将文件从本地机复制到远程计算机上 (Upload)。FTP 是获取免费软件、共享软件资源不可缺少的工具。

通过 FTP 可以获得的信息资源类型非常广泛。广义地说, 任何以计算机方式存储的信息均可保存在 FTP 服务器中。可以 FTP 形式获得的信息资源有: 一些书籍的电子版, 如《圣经》、《莎士比亚全集》等; 电子期刊与杂志; 某些政府机构发布的信息, 如美国国家统计局的统计数字; 美国航空航天局某些实验的信息、最新图像文件、太空探测照片等。另外, 就是大量的免费软件和共享软件 (可先试用, 再注册交纳一定费用后就可获得软件的使用权), 包括 DOS, Windows, Macintosh, UNIX 下的软件以及 Apple, Novell 及其他主要厂商产品的最新资料, 涉及各类应用程序、专用计算软件、汉字平台、文件

压缩工具、代码转换工具、游戏软件、反病毒程序、通信软件等许多种类型。目前,国内各高校、科学院、机关团体的公共 FTP 站点和内部 FTP 站点上都有丰富的 FTP 资源。

如同其他的互联网服务一样,FTP 采用客户机-服务器 (Client/Server) 结构,用户在使用 FTP 服务时,需要输入远程主机的用户名和口令才被允许登录。在因特网上的许多 FTP 站点提供匿名文件传输服务 (Anonymous FTP)。用户登录时,以 Anonymous 作为用户名。常用的 FTP 软件有 LeapFTP, CuteFTP, Serv-U 等。

通过 FTP 获取文件的一般操作步骤为:

(1) 登录。即为用户建立与某 FTP 服务器的连接,可在地址栏中输入要登录的主机地址 `ftp<ftp 主机地址>`,启动 FTP 软件程序。在实现了与远程 FTP 主机的连接后,需输入用户名和口令。若访问的是匿名 FTP,则不需要账号和口令,以“anonymous”作为用户名,E-mail 地址或“guest”作为口令;一般这种访问被限制在公共目录 (public 或 pub) 下。

(2) 浏览目录,找到需要下载的文件。

(3) 选择保存文件的位置。

(4) 下载文件。将文件从源地址拖动到目的地址,即完成了文件的复制。有时在下载前还要设置文件传输参数,选择文件的传输模式。如:ASCII 模式——用于传输文本文件,二进制 (Binary) 模式——用于传输非文本文件。

上传文件的方法与上述步骤类似,只需将文件从本地拖动复制到目的地址空间即可。

4. 用户服务组信息资源

各种各样的用户通信或服务组是互联网上最受欢迎的信息交流形式,包括新闻组 (Usenet Newsgroup)、邮件列表 (Mailing List)、专题讨论组 (Discussion Group)、兴趣组 (Interest Group)、辩论会 (Conference) 等。虽名称各异,但功能较为相似,都是由一组对某一特定主题有共同兴趣的网络用户组成的电子论坛。在这些论坛中所交流的文章即是一封封的 E-mail,因此可将其理解为电子邮件功能的进一步扩展,使人们能更便捷地进行多向交流。迄今为止,你可能想到的任何一个主题均有与之相对应的组 (Groups),总量难以估计,参与者为全世界的网络用户。

以上述各种用户通信组形式传递和交流的信息就构成了互联网上最流行的一种信息资源。主要包括某个学科领域的新闻、研究动向、最新成果发布、交谈、质疑解惑、讨论、评论等,它是一种最丰富、最自由、最具有开放性的资源。其信息交流的广泛性、直接性、专指性与其他信息资源类型所不能比拟的。在这里你既可以与遍及世界各地的各个领域的专家直接交流和讨论,更新知识,澄清和发展自己的观点,也可以就生活中的一些问题、小窍门等求教于一些热心助人的朋友。对于研究人员来讲,应该将其视为获取出版渠道以外的研究性信息并与业界同行取得交流的一个重要渠道。可以用于了解学术动态、参与讨论,还可以活跃自己的创新思路,或就某些疑难问题或技术环节求助于同行的提示、参考意见,分享有益的经验,其直接、方便、快捷和非正式等特点都对学术研究大有裨益。

使用 Usenet, 要通过一个特别的新闻阅读程序 (又叫新闻阅读器, Newsreader)。在 UNIX 系统中, 有四种阅读器: nn, rn, tin 和 trn。而在 Windows 系统中, 也有多种不同的新闻阅读器可供选择, 如 WinTrumpet。现在也可以将 WWW 浏览器作为新闻阅读器来阅读新闻组消息, 可在浏览器地址框中输入 “news:<新闻组名称>” 来连接相应的新闻组服务器。

而邮件列表 (Mailing List) 则可使用任何一种电子邮件系统来阅读新闻和邮件, 并允许用户向能够做出响应的人发送邮件。它实际为参加者的一份通讯录, 每个参加者均可向这个群组发送邮件, Mailing List 再转发给组内所有的参加者, 实现一对多的交互。日常生活中, 任何有某种共同需要的人群都可组成一个 Mailing List, 如: 校友联谊会或学习某课程的班级的学生均可创办一个 Mailing List, 作为教师与学生、学生与学生之间日常交流、答疑、讨论的园地, Mailing List 是应用网络辅助教学的一种重要形式。目前, 有很多学科专业的电子期刊、行业快报等以此形式发行、传播。

上述用户服务组通常都具有以下几种功能:

- 订阅该组;
- 取消订阅该组;
- 读取组中文章;
- 发布文章至该组;
- 跟随文章 (在读取文章中加入心得或建议);
- 回信;
- 转信 (将所读取的文章转发给其他人)。

5. Gopher 信息资源

在 Web 资源出现前, Gopher 是互联网上最流行的分布式信息资源体系, 曾经是最为广泛而又功能强大的系统。它是一种基于菜单的网络服务, 能为用户提供广泛、丰富的信息, 并允许用户以一种简单、一致的方式快速找到并访问所需的网络资源。全部操作是在一级级菜单的指引下, 用户只需在菜单中选择项目和浏览相关内容, 就完成了对互联网上远程联机信息系统的访问, 而无需知道信息的存放位置和掌握有关的操作命令。这也反映了 Gopher 的另一优势, 即它可以跨越多个计算机系统, 仅通过运行本地计算机的 Gopher 客户程序就可以与世界各地任何一个 Gopher 服务器连接并共享信息。此外, Gopher 还可以提供与其他信息系统, 如 WWW, FTP, Telnet, WAIS 和 Archie 等的连接。

互联网上曾有几千个 Gopher 服务器, 存储有各种各样的信息。许多 Gopher 服务器分布在大学、公司或其他组织机构内, 每个 Gopher 服务器内都存放着本部门或本地区用户感兴趣的信息。一般学校 Gopher 服务器所提供的信息资源有:

- 资源目录或索引——该服务器所提供的全部资源的索引和介绍。
- 图书馆——本大学图书馆目录及电子文献 (文学名著、电子期刊或学位论文等)。
- 数据库——本服务器所提供的一些公共或专用数据库。
- 校内院系资料——包括概况、学位、课程、师资、设备和招生信息等。

- 校内活动信息——包括报告、演讲、学生社团、学会活动和体育比赛等。
- 地方信息——如关于当地交通、生活、风景和习俗等的介绍。
- 生活信息——如天气预报、饮食服务和租住信息等。
- 其他——通向其他 Gopher 服务器的连接。

Gopher 曾经以其简单、统一的界面，方便、易用的特点和丰富的资源构成了互联网上一种重要的资源类型。某些专业人员还认为：从总体上讲，Gopher 信息质量要优于 Web 资源，因为在 Gopher 信息被列入菜单目录等生成过程中要经过一定的审核，对其质量有一定的监控。然而随着网络的发展，只能提供文本信息的 Gopher 如今已成了昨日黄花，芳踪难觅，许多 Gopher 服务器都已被 Web 服务器所取代。

1.3 互联网信息资源评价

1. 评价的意义

与传统的信息资源相比，互联网信息资源的首要特点即广泛的可存取性（Accessibility）。网络信息的自由存取和易用性，导致了网络信息资源的繁盛。任何可想像到的学科、主题领域均有大量的网络信息产生。网上信息广泛、丰富，却缺乏组织和质量控制，呈现着无限、无序、优劣混杂的发展状态。正如国外学者指出的那样：“网络的最大优势，同时也是其最薄弱之处是任何人、任何机构随时都可以在互联网上发布信息，生产出一个数字化的知识数据单元 DLO，即类似文献的电子文件（Document-Like-Object）。网络用户均明显地感觉到由信息过载（Information Overloaded）所引起的困惑和无所适从。用户对获取高质量网络信息的期望值日益提高，他们希望所获取的网络信息是有效的、可靠的、权威的、相关的和适用的。面对网络信息资源的质量不均衡状态，发展网络信息资源的评价方法及相关标准和工具是至关重要的。进行网络信息资源评价，从信息海洋中经过甄别，挑选出有学术价值或利用价值的精华部分，推荐给网络用户使用，可以较好地屏蔽一些信息污染或检索噪音，大大地提高用户利用互联网信息资源的效率。

另外，互联网改变了传统的信息发布和评价程序。在印刷型文献时代，信息的评价和过滤是由编辑、出版者、评论人员或权威部门（如政府机构等）、学科专家等来完成的；而在互联网时代，发表自由使信息的生产跳过了编辑出版这一至关重要的质量控制环节，对网络信息的评价更多的要由网络用户自己承担。因此，网络用户必须了解和掌握一些对网络信息资源的评价标准和评价方法，以使自己的上网“冲浪”是建立在某种分析和判断基础上的“智能性”的访问，同时还能对自己所搜集和获取到的信息资源的价值有所判断。而网络用户同时也会成为网络信息的编辑、生产者，掌握有关的评价标准，能够使其在编辑和生产网络信息的过程中有所规范，以便在几个关键的方面注意按照标准的要求去做，进而提高和改善网络信息的质量。

2. 网络信息资源评价的主体与客体

网络信息资源评价的主体通常包括：以赢利为目的提供网络资源评价服务的商业机构；以研究为目的开发网络信息服务资源的非营利性组织；以教学、科研服务为主，利用自身资源、人力优势开展评价研究服务的高等院校系、所、图书馆；有兴趣、有能力、目的不一的个人。由于对网络信息资源进行评价的目的不同，导致各类评价主体在评价的标准、范围、方式、方法和侧重点等方面各有不同，评价结果差异较大。

网络信息资源评价的客体通常包括网站和网页。具体地说，包括网站和网页所承载的信息内容及承载网络信息内容的载体网站和网页。我们不仅要评价网站中具体信息的内在属性和外在属性，同时要关注发布这些信息的网站的组织设计水平、运作方式及人员素质等与信息资源质量密切相关的因素。

3. 网络信息资源评价的方法与步骤

网络信息资源评价的方法包括定性评价方法和定量评价方法。

(1) 定性评价方法：即按照一定的评价标准，对所评价网站、网页的内容与形式特性进行评价。这是目前网络信息资源评价的主要方法。就实施评价的主体区分，目前较常见的是用户评价法和第三方评价法（第三方即相对于信息资源发布者与信息用户而言的第三者）；就实施评价的主要方式区分，包括专家评价法、问卷调查及统计法等。定性评价方法的核心在于选择科学合理的评价指标体系，这关系到评价的客观性、公正性、科学性和合理性。由于目前尚未形成统一、科学的评价指标体系，各类评价往往不可避免地带有浓厚的主观色彩，从而影响到评价结果的客观性、公正性和可信度。

(2) 定量评价方法：即按网络信息资源自身特征和规律，采用数量分析方法进行客观评价。这是目前正努力研究探讨的网络资源评价方法。如对网络信息资源被访问、登录、链接和检索次数的统计分析、借鉴传统引文分析法评价网络信息资源价值和重要性的网络计量法都属此范畴。定量评价方法可用于评价分析网络信息资源的分布情况和被利用情况。从理论上说，定量评价方法具有较强的客观性和公正性，但由于网络信息资源具有高度离散、动态、无序的特点，在目前的技术条件下，进行完全的量化评价尚有一定的难度。

鉴于定性评价方法和定量评价方法各自的特点，在现阶段将定性和定量评价方法进行有机结合，无疑具有相当的现实意义。基于层次分析法的网络信息资源评价方法正是两者融合的产物。

层次分析法是一种综合人的主观判断来分析复杂问题的拟定量方法，适用于解决结构复杂，准则繁多又不易完全量化的现实问题。它通过找出研究对象相关的各主要元素，分析各元素间相互关联与隶属关系，构建等级层次结构模型。对同一层次的各元素针对上一层次特定元素的相对重要性进行两两比较，建立判断矩阵，计算被比较元素的相对权重，并进行一致性检验；进而计算各层次相对于研究对象的合成权重，进行总排序，得出分析结果。运用层次分析法，建立网络信息资源评价的层次结构模型和评价体系在现阶段不失为一种有效的、切实可行的方法。

网络信息资源评价的步骤为：（1）确定网络信息资源评价的对象及范围。（2）确定网络信息评价指标，并收集评价对象的有关数据与信息，参照指标体系对比、分析、打分予以评价。（3）为评价结果排序，并写出评价报告，供用户参考。（4）根据用户的反馈信息，对评价方法、过程等进行调整、核实和归档。

4. 评价的标准

科学合理的指标体系是网络信息资源评价的基础和关键。近年来，国内外学者对此均做了不少有益的探讨和研究，提出了各种标准和原则，目前仍处在探索阶段，尚未形成一套公认的方法、准则和评价标准。综合各方专家学者的研究成果，大致可勾勒出网络信息资源评价指标体系的框架。

1) 网络信息资源内容评价指标

网络信息资源的内容是决定信息资源质量与价值的最重要因素，有关内容的评价指标包括：

- 准确性——这是信息资源内容质量最基本、最重要的评价指标。包括信息内容是否清晰地显示所依赖的数据与依据，是否来自权威机构或公开合法的文献，事实和数据是否经得起推敲、验证，遣词造句是否缜密、流畅、准确、清晰。评价时可利用一定的专业知识加以评鉴或与其他相关资源进行比较。
- 权威性——指网络信息资源是否来源于有声望的专家、学者或组织、机构、团体。信息制造者不同的身份和造诣往往产生不同质量的信息内容。评价时可通过对作者和机构的了解、对信息内容来源的了解加以判断。
- 新颖性——指网络信息资源的更新频率和周期。评价时可可通过查看文档的创建、发布、修改、更新日期加以判断。
- 独特性——指网络信息资源内容的特色及风格。有特色才有生命力，才能生存。评价时可可通过了解网络信息资源的内容、观点、方法手段、对象群、收录范围是否有独到之处及是否可从其他渠道获取入手。
- 涵盖度——指网络信息资源内容涵盖的广度与深度。评价时可可通过了解信息资源主体的详尽程度和涉及相关领域的范围加以判断。同时，是否提供数量可观的相关资源链接也是反映涵盖度的一个重要方面。
- 客观性——指网络信息资源内容的客观性与公正性。评价时可从信息内容与作者或团体的利益是否有关联、内容的阐述是否理智、中立地反映各方思想和观点等方面入手。

2) 网络信息资源网站/网页的组织与设计评价指标

任何网站/网页都是内容与形式的统一体，两者相辅相成，交相映衬。因此在确立网络信息资源内容评价指标的同时，不能忽略内容赖以存在的形式与载体的评价指标体系的建立。

- 目的指标——指网站/网页设立的目标、确定的用户群、提供信息服务的层次与程度等网站建设的最基本因素。评价时可从网站声明中了解信息并判断。