



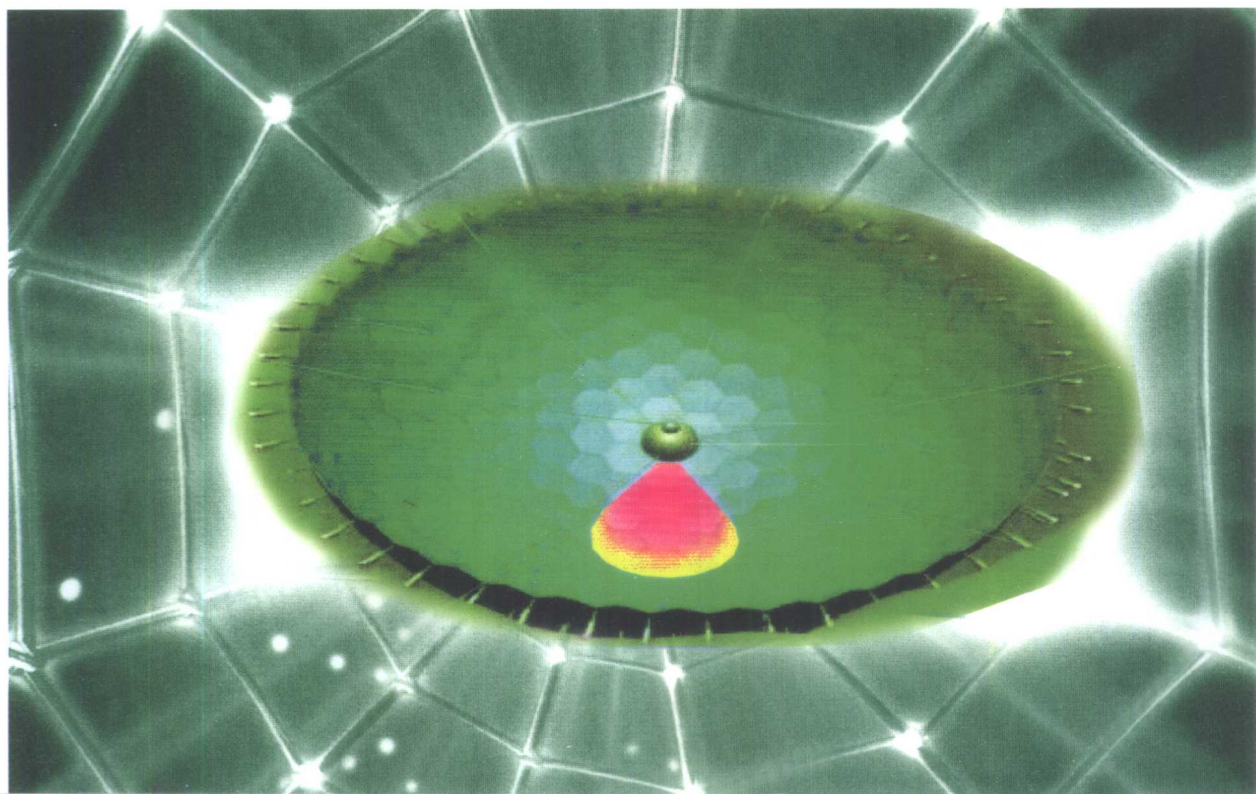
面向21世纪课程教材

信息管理与信息系统专业教材系列

因特网信息资源检索与利用

符绍宏 雷菊霞 饶伟红 编著

清华大学出版社



面向 21 世纪课程教材
信息管理与信息系统专业教材系列

因特网信息资源检索与利用

符绍宏 雷菊霞 饶伟红 编著

清华大学出版社

(京)新登字 158 号

内 容 简 介

因特网信息的无限、无序、优劣混杂,缺乏统一的组织与控制等特点给人们有目的的查找和利用信息造成了很大的不方便。在因特网这个浩瀚、动荡的信息海洋中,准确、及时、有效地查找并获取所需要的信息,对所有网络用户来说都是十分重要的,同时也非常具有挑战意味。本书面向广大因特网用户,针对利用网络检索与获取对用户研究开发、工作学习有价值的网络信息的需求,系统地介绍因特网信息资源的特点、种类、信息利用价值及评价、鉴选标准;介绍因特网信息检索的基础知识,各类型常用网络信息检索工具的特点、比较、评价、使用和检索方法;详细介绍了学术性、研究性及各类专门信息的网上检索与查找方法、途径及重要站点。本书对培养网络信息查找与利用能力,提高在网络信息环境下的生存和发展水平,都具有实用价值。

本书可作为高等院校网络信息检索课程的教材,还可作为广大网络用户的实用参考书。

书 名: 因特网信息资源检索与利用

作 者: 符绍宏 雷菊霞 饶伟红 编著

出 版 者: 清华大学出版社(北京清华大学学研大厦,邮编 100084)

[http:// www. tup. tsinghua. edu. cn](http://www.tup.tsinghua.edu.cn)

责任编辑: 魏荣桥

印 刷 者: 北京牛山世兴印刷厂

发 行 者: 新华书店总店北京发行所

开 本: 787×960 1/16 **印 张:** 15 **字 数:** 317 千字

版 次: 2000 年 6 月第 1 版 2001 年 7 月第 2 次印刷

书 号: ISBN 7-302-03822-8/F·267

印 数: 5001~8000

定 价: 19.50 元

序 言

人类正迈向知识经济的新世纪,为了使我国高等教育能适应国内外急剧变化发展的形势,1995年春,原国家教委正式推出和开始实施了“高等教育面向21世纪教学内容和课程体系改革计划(以下简称教学内容改革计划)”。改革的总目标是:“转变教育思想,更新教育观念,改革人才培养模式,实现教学内容、课程体系、教学方法和手段的现代化,形成和建立有中国特色的社会主义高等教育的教学内容与课程体系,提高教学质量,培养适应21世纪需要的社会主义现代化建设者和接班人”。近期目标是:“用5年左右时间,集中优势力量,开展集体攻关,努力形成一批优秀的研究和改革成果,使我国高等教育的教学内容和课程体系相对落后于科技、经济、社会发展的状况有较大改观,为21世纪初叶大范围提高我国高等教育质量打下良好基础”。在高等教育改革的系统工程中,教学改革是核心,而其中教学内容和课程体系的改革,又因其直接反映教育目的和培养目的,是培养人才素质、提高教育质量的核心环节,故又是教学改革的重点和难点,是深层次的教学改革。为了反映和固化各专业基础课和主干课教学内容和课程体系改革的成果,教学内容改革计划还提出要编写出版一批高水平和高质量的“面向21世纪课程教材”。

作为国家教委级221个大的立项项目之一(理科27个之一),“信息管理类专业教学内容和课程体系改革(以下简称信息管理专业教学改革研究)”,也于当年立项并开始启动。项目由北京师范大学牵头,有中国科学技术信息研究所、复旦大学、北京大学、武汉大学、中山大学、四川大学、山东大学、天津师范大学、北京联合大学和吉林工业大学等11所高校和科研院所参加。4年来,经过参加人员和广大教师积极认真的改革研究和实验,取得了重要进展,明显提高了实验班级毕业生培养质量的实际效果,并编写出和推荐了头一批“信息管理与信息系统”专业教材,本书就是这批书中的一种。

以下几点是“信息管理专业教学改革研究”项目考虑教学内容和课程体系改革问题的基本出发点,也是编写这套教材要努力做到的。学生是教学工作的主体,使用这套书的学生,在学习每门课时了解整体学术背景,无论是对加深当前学习内容的理解,还是对认识和发挥自己“主体”地位的作用都有必要。

1. 关于“信息管理”和“信息管理学”

“信息管理”是一个范围很宽,正在发展的概念。按项目组目前研究的认识,“信息管理”主要是指信息资源的管理,包括微观上对信息内容的管理——信息组织、检索、加工、服务等,及宏观上对信息机构和信息系统的管理。“信息管理学”是一门研究人类信息管理

活动的规律及应用的学科,它以数学、管理科学、信息科学与技术作为基础,涉及多学科(理、工、农……)及多领域(经济、管理、法律……)的综合性交叉学科。高附加值的信息是知识创新的关键因素,但浩如烟海的信息,只有经过有效获取、科学加工和有序管理,才能成为可利用的资源,信息管理学所研究和提供的就是这种理论知识和技术。在国内外经济知识化、社会信息化迅速发展和激烈竞争的形势下,信息管理、信息科学、信息技术、信息产业与信息教育共同成为信息社会战略上需要优先发展的行业。有人把网络比作信息高速公路,计算机软硬件比作在高速公路上跑的“车”,信息资源比作“货”,当前的问题是“车”严重空载和“货”的质量不高,信息管理就是“加工生产”大量优质信息产品的知识工具,而问题的症结就在于需要培养出大批高素质的信息管理人才。

2. “信息管理学”专业的基本内涵

什么是决定“信息管理学”专业核心教学内容和课程体系的基本内涵?应当说目前还没有明确的界定,但根据学科的发展和社会定位,根据信息科技发展对其产生的影响,根据培养目标的定位和国内外许多学者的建议,“信息管理学”的专业内涵应建立起一个由上到下的信息管理学的学科体系。如 Wilson 的观点:“应从信息的产生,信息的组织与处理,信息的传播与利用来定义”。有的学者还提出,一个完整的信息管理科学的知识体系应该包括信息的上、中、下“三游”。而作为一信息管理和服务的专业人员,应该要“了解信息的上游,掌握信息的中游,提供下游的服务”。

所谓信息的上游,乃是了解信息的本质、信息的特征、信息的种类、信息产生的过程、信息生产的主要机构及信息如何传播等。而这里所谓的信息,不应只限制在学术信息,也不应只强调图书媒体,而应该包括所有相关的信息媒体,及所有可能的相关信息的研究。

所谓信息的中游,则包括信息的搜集、信息的整理、信息的组织与分析、网络环境下多媒体信息的描述与组织、信息系统、信息的储存、信息的检索、信息的深加工等。

而下游的信息利用与服务,则是了解我们的使用者,了解他们的信息需求,然后提供他们所需的信息服务。

如果我们依照信息管理人员的工作内容来看,信息上游的了解是所有信息管理人员所应必备的知识。这个基本常识掌握后,才能因工作性质的不同,调整应掌握的知识面。

信息的组织与分析应是本专业最核心的专业知识,本专业所有其他的专业知识都应该是由此而发展,这也是我们跟别的专业有所区别的地方。

由以上分析是否可认为本专业的基本内涵是:信息的产生和传播,信息的组织分析,信息的利用与服务。以上认识,也是我们提出分层次的专业基础课和核心课内容体系的基本出发点。

3. 转变观念,建立新的培养模式

为满足对信息管理人才的需要,在新修订的我国高等教育本科专业目录中,在原科技

情报、信息学、经济信息、管理信息系统等专业基础上,调整合并设立了新的“信息管理与信息系统”专业。1999年全国有150多所高校在这个专业招生,成为今年理科最热门的专业之一。如何办好这个新调整的专业,提高培养质量?关键在于**转变教育观念,改革培养模式**。信息管理是科技含量高、知识更新快、交叉学科多、国际化程度高、应用范围广、服务性强的发展中的新学科,要求本科培养的人才具有广泛和多层次的适应性,为此必须改变以适合在某一种行业或领域从业为培养目标的旧观念,而把拓宽加深基础和提高以创新能力和实践能力为主的全面素质作为培养目标;改变过去重理论、轻实践、多灌输、少自学的旧教学模式,而把充分开发潜能,使学生做到“自奋其力,自致其知”,作为提高人才素质的运作机制,突破传统基础观念对“加强基础”的理解,不仅要通过系统的理论知识学习打下学生聚合思维的基础,还要通过教学和实践环节的系统改革,培养学生对科学问题的兴趣和锲而不舍的探索精神,打下创造能力所需要的发散思维和人格品质的基础,并要将这两种思维进行整合,内化成学生的认知结构。如开设将信息的采集、描述、组织、存储、检索、分析和利用等若干知识点互相结合应用,充分体现自主学习的“课程设计”课;在教学计划中安排整个第八学期进行面向科研、应用开发和服务等实际问题的毕业实践(毕业设计和毕业论文)等。教学改革是一项系统工程,按系统工程原理,局部和可能大于也可能小于整体,各门课的改革和教材编写要从整体出发考虑,才能实现系统总体目标的优化。

4. 信息科技的发展和教学内容的改革

信息对知识经济发展的意义,可分为两方面,一方面是信息成为知识创新的因素;另一方面,为了有效获取和利用信息,要促使信息科技的发展。特别是90年代以来,网络科技的迅猛发展,推动了知识经济的加速发展。这使得知识经济和信息科技成为大学教育面临的社会背景和重要技术背景,这无疑对信息管理学科内容和体系的变化首当其冲地产生重大影响。因此,信息管理专业教学内容和课程体系改革的重点和难点,不仅要重视内容和体系改革的整体优化,精简陈旧内容,加强信息科技内容的比重,给学生更多实践、研讨和自主学习的时间,积极应用信息技术于教学手段和方法,更重要的是专业教学内容的定量化和信息化,在信息的产生传播、信息的组织分析、信息的利用与服务三个层面的专业教学内容中,以信息科技和相关科技的发展为融合点和生长点,拓宽理论基础,改造传统教学内容,改进研究方法,改变“两层皮”的状况,做到有机结合,形成严谨的专业内涵,构造新的内容和体系。这样才能从根本上开发学生的潜能,培养出面向未来和面向世界的高质量信息管理人才。

5. 信息素养教育

信息科技,特别是网络科技的迅猛发展,也导致人类的沟通与信息交换方式由过去的人际互动(human-human interaction)模式,变为以人机互动(human-machine interaction)为主的模式。终身学习、如何学习、能力导向学习和开放学习成为新的教育理

念的重要内涵。为满足知识创新和终身学习的需要,发达国家纷纷将信息素养(information literacy)教育,作为培养 21 世纪人才能力的重要内容。所谓“信息素养”或“信息能力”,是指使用计算机和信息技术高效获取、正确评价和善于利用信息的能力。目前美国全国从小学、中学到社区学院和各大学都已全面将信息素养纳入正式的课程设置之中。“信息素养”是一个有时代性的、重要的教育议题,是未来信息社会考察一国人力素质和生产力的重要指标。作为信息管理专业的学生,不仅比其他专业更有必要提高自己的信息素养,还有责任在对大众实施信息素养教育计划的重任中发挥作用。

6. 教材特点

首批推荐的这一套教材,反映了经过该项目组成员和广大教师的努力而取得的改革成果。具有新的教育观念和内容,作者水平较高,校际合作、老中青合作范围较大,适应面较广等特点。

本套书的作者都是国内信息管理方面的专家或后起之秀,本套书不仅是他(她)们多年教学改革及经验的总结,也是他(她)们最新科研成果的结晶。有的教材出版印刷过几次,获过奖,此次再版增加了新的研究成果;有几位青年作者近年刚从国外深造归来,在他(她)们的书中反映了学科前沿的最新进展。

经过专业目录调整,全国有多所高校设立了信息管理与信息系统专业,尽管原有专业背景不同,但改革的方向和目标将逐步融合。此外,国家自考大专和自考大学本科也新增加了信息管理与服务专业,不少院校的夜大和函授大学也有此专业,还由于信息管理是发展较快的新兴学科,故可估计此套书出版后的读者群将比较大,尽可能做到兼顾不同层次的读者,也是编写这套教材时重视的一个问题。

一本好的教材一定要经过多次修订才能成熟起来。这套书中有的是第一次正式出版,尽管编著者和出版社为此付出了很大努力,但难免有不足之处,欢迎读者批评指正。

这批教材能列入“面向 21 世纪课程教材”及时出版,要衷心感谢教育部高教司理工处领导的大力支持,也要衷心感谢高等教育出版社、清华大学出版社、武汉大学出版社编辑同志的积极配合与大力支持。

康仲远

于北京师范大学

1999 年 9 月 10 日

前 言

近年来,因特网的广泛应用和发展构成了人类信息环境变化的重要层面。因特网使世界范围内的信息交流、资源共享成为可能,从而大大拓展了人类的信息空间,同时它也对传统的信息组织、检索和获取方法构成了很大冲击。网络信息的无限、无序、优劣混杂,缺乏统一的组织与控制等特点给人们有目的的查找和利用信息造成了很大的不方便。上网用户在面对大量纷繁复杂的信息、数据时,都会感觉到由信息过载(information overloaded)引发的困惑和茫然。在网络信息世界这个浩瀚、动荡的信息海洋中,准确、及时、有效地找到、获取与自身信息需求相关、切题和适用的信息,对所有网络用户来说都是十分重要的,同时也非常具有挑战意味。

随着网络信息组织与检索实践的不断丰富和发展,各种类型、各种品牌的网络信息检索工具如雨后春笋般不断涌现,其各自的检索功能与服务在不断完善;结合网络信息资源检索特点的各种行之有效的检索策略、方法与技巧等也在逐步的探究与形成中。尽管因特网上的信息资源浩如烟海,没有全面组织、没有综合索引,而借助相应的检索工具,遵循有关的查找方法,仍然可以“取之有道”。

网络环境下的信息检索与传统信息环境下的检索有很大不同,网络信息检索所具有的多样性、灵活性也远远超出了传统的信息检索。网络信息检索(NIR——networked information retrieval)将不仅代表着在网络信息环境下发展起来的一种新型检索模式,而且将发展成为信息检索领域一个重要的学科分支。随着网络信息检索实践的日渐丰富和发展,结合网络信息资源的特点,研究网络信息检索的规律、方法,以及网络信息检索系统的设计、开发思想和技术,探究进一步提高检索效率的途径和方法等,亦即探索“信息检索”这一学科在新的信息环境下的拓展和应用具有十分重要的意义。

本书面向广大因特网用户,针对利用网络检索和获取对其研究开发、工作学习有价值的网络信息资源的迫切要求,系统介绍因特网信息资源的特点、种类及信息利用价值,和对因特网信息资源的评价、鉴选标准;网络信息检索的基础知识;各类型网络信息检索工具的特点、比较、评价、使用和检索方法,特别针对学术性、研究性信息的网上检索与获取介绍有关的查找方法、途径及重要站点。本书着重讨论因特网信息检索,是在读者对因特网已有基本了解,掌握其基本应用的前提下,着眼于其更高层次、更深入的信息需求并给予指导,对培养读者的网络信息查找与利用能力,提高其在网络信息环境下的生存和发展水平,具有很强的实用价值。

本书基于编者近年来的学习、研究与教学实践,其初稿曾作为北师大相关课程和专

题讲座的讲义,在几轮教学中反应良好。但由于网络信息检索研究是一个较新的领域,尚无较成熟、系统的理论阐述或概括,加之本书编者的水平、视野有限,书中的偏颇、谬误之处在所难免。另外,尽管在编写过程中我们注重加强了内容的新颖性,但因特网极富变化的特征仍可能使本书的某些内容在出版时就有些过时了。上述不足之处均请读者在使用中多多包涵并不吝赐教。

本书分工如下:第1~2章由符绍宏编写;第3~6章由雷菊霞编写;第7章由饶伟红编写;第8章由符绍宏编写;第9章由饶伟红、符绍宏编写。

本书写作过程中,得到康仲远教授、袁名敦教授的大力支持与帮助,在此表示衷心感谢。

编 者

目 录

第 1 章 因特网信息资源	1
1.1 因特网上的信息资源	1
1.2 因特网上信息资源的特点及信息利用价值	2
1.3 因特网上信息资源的种类	3
1.4 因特网信息资源的评价	8
第 2 章 因特网信息检索概论	15
2.1 网络信息检索的发展	15
2.2 网络信息检索的一般方法	15
2.3 网络信息检索的特点	18
2.4 网络信息检索工具	19
2.5 对应客户机/服务器模式的网络信息检索标准 Z39.50	23
第 3 章 网络信息检索工具(一)——搜索引擎	27
3.1 搜索引擎的原理	27
3.2 搜索引擎的检索功能	28
3.3 检索结果的显示	30
3.4 搜索引擎介绍	31
第 4 章 网络信息检索工具(二)——网络资源目录(Web Directory)	49
4.1 目录型网络检索工具原理	49
4.2 网络目录与搜索引擎的比较	51
4.3 网络目录介绍	52
第 5 章 网络信息检索工具(三)——多元搜索引擎	64
5.1 多元搜索引擎	64
5.2 多元搜索引擎的检索特性	65
5.3 多元搜索引擎介绍	66
第 6 章 网络信息检索工具(四)——专门信息检索工具	81
6.1 查询个人地址、电话、电子邮件及相关信息	81
6.2 查询地图及相关信息	87

6.3 查询多媒体信息(图像、声像)	91
第7章 其他网络信息检索工具	98
7.1 FTP 资源检索工具	98
7.2 Mailing list 邮件列表检索工具	108
7.3 远程登录 Telnet 检索工具	118
7.4 用户网 Usenet 检索工具	122
7.5 专业性网络信息检索工具	127
第8章 学术信息的网上检索与获取	157
8.1 OCLC 的 FirstSearch 检索系统及服务	157
8.2 CARL 的 Uncover 检索系统及服务	164
8.3 图书信息的网上检索	167
8.4 期刊信息的网上检索	172
8.5 会议信息及会议文献的网上检索	179
8.6 学位论文的网上检索	181
8.7 专利的网上检索	182
8.8 标准信息的网上检索	186
8.9 国际组织、政府机构及出版物信息的网上检索	189
8.10 学术文献的网上获取	197
第9章 网络信息资源管理	199
9.1 网络信息检索研究	199
9.2 网络信息资源的组织与管理	209
9.3 网络环境下的版权保护	217
9.4 网络信息检索的未来发展	220
参考文献	225

第 1 章 因特网信息资源

1.1 因特网上的信息资源

因特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。在以因特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都是一项挑战。而要进行有效的网络信息检索,首先就必须对因特网上信息资源的发布、种类和利用价值等有较全面的认识 and 了解。

因特网这个庞大的信息宝库其实并非如某些人想象的那样应有尽有、唾手可得。而围绕着因特网的信息资源与服务,确实存在着许多误解,如:

- 认为因特网无所不包、无所不能;所有信息均可在网上找到;
- 认为在因特网上找到的、下载下来的信息都是正确的;
- 认为因特网可取代其他媒体或信息渠道;
- 认为因特网上的信息无太多利用价值;仅限于聊天、菜谱、各种发烧友俱乐部等非正式、娱乐性信息,缺乏严肃类信息。

作为检索和利用因特网信息资源综合策略的第一步,就是要了解网上有什么和不可能有什么信息内容,进而明确因特网在整个信息交流圈、信息链中所处的位置,知道什么样的问题可以在因特网上找到答案。从而在面对各种媒体、各种信息检索系统、检索工具时,能够根据检索需求的不同,做出正确的选择。在当今的信息环境下,能够“在适宜的时机,以适宜的形式、适宜的价格找到确切的信息”是具有较高检索智能的重要体现。

1.1.1 网上有什么样的信息内容?

(1) 政府信息

指国际组织、各国政府及其相关部门所发布的信息。包括:有关组织机构的宗旨、业务范围、人员、出版物、最新消息发布;还有各种法律、法规或相关政策信息;从统计数据到被延迟的立法等,内容相当广泛。政府信息可以说是网上最有价值的一部分信息,由于出自官方,所以具有权威、可靠等特点。另一个突出的优点是价廉,其中大部分是免费信息,可以免费检索、获得,甚至有许多信息还没有版权限制。

(2) 科研信息

主要指各专业学术机构(如:学会、协会、研究所等)所设立的网站及其相关信息。包括:目标、宗旨、成员、主要出版物、最新学术活动安排(如:会议、展览等)、各种学术性电子期刊、学科专题论坛、讨论组,各学科应用的工具性计算机软件和研究性计算机软件等。

(3) 教育信息

主要有各大学所设立的网站及相关信息。其内容相当广泛,有该校各学院、系、专业的介绍,学位、奖学金的设立,入学申请表、校历,以及学校周边环境、生活设施、公共交通。还有各学科专业的教学计划、课程表,及教师的个人网页(通常含其个人资历、所任课程的教学大纲、研究成果、著述、近期研究活动等)。另外,因特网已经发展成为一种重要的教育手段、学习环境,它所包含的教育信息是极为丰富的,如:各种远程教育的课件;用于学生自我学习的交互式虚拟教室;还有大量为教师教学提供参考的或素材的信息,如科学、艺术和人类感知的各种图片、照片等多媒体网站及数据库。

(4) 文化信息

主要包括各类信息媒体(如:报纸、杂志、电视、广播等)的网站,和世界各地图书馆的数字化馆藏及其联机公共检索目录(OPAC—online public access catalog)。

(5) 消闲娱乐性信息

被称为是网上“最成功”的领域。包括:旅游信息,各旅游城市的风景名胜、特产、民俗,以及宾馆、饭店、交通等信息均可在网上找到。购物信息,各种网上书店、网上商场的商品、价格、销售信息。还有为任何一种可想象出的兴趣、爱好(如:足球、音乐、电影、收藏、烹饪等)而开设的网站、讨论组、网络俱乐部等。

1.1.2 网上不可能有什么样的信息?

尽管网上信息发展很繁盛,还是有许多重要信息是不可能在网上公开,或使用户自由访问、获取的。如:有关企业、公司的战略、发展规划,商业、贸易秘密;商业性数据库和大多数有版权的印刷物。(一般来说,有版权的印刷资料的拷贝或其他商品不能在网上得到,但因特网是不断变化的,情形有时会改变。如:近来有许多有版权的图书、期刊等出版物在因特网上部分地公开其内容,供用户自由浏览;某些数据库也把某一时间段的数据库公开,供用户免费检索。(如:著名的教育文献数据库 ERIC 和医学文献检索数据库 MEDLINE。))

1.2 因特网上信息资源的特点及信息利用价值

因特网作为数字化、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。什么是因特网信息的独特之处？它具有什么优势和局限性？因特网信息具有哪些重要的信息利用价值？

因特网信息的独特之处可以概括为：

- ① 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。
- ② 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、图象、声音、动画等。
- ③ 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、消亡等都时有发生，难以控制。

总之，因特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、快速和交互性的特点。因特网信息在以下一些方面更具有优越性：

- ① 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。
- ② 新颖、深入。因特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、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，或是因为其内容太新或太专而未被纳入正式文献交流渠道。
- ③ 广泛、直接交流。因特网扩大了人际交流的范围，提供了更多的直接交流机会。如参加友思网(Usenet)的新闻组(newsgroups)、讨论组(discussion group)、邮件列表(mailing list)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。
- ④ 非正式和自由发表园地。因特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，因特网信息资源不是传统信息资源的复制，因特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，不掌握因特网这一信息工具，不利用因特网信息资源，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。

1.3 因特网上信息资源的种类

因特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，可将因特网信息资源划分为以下几种类型：

1.3.1 万维网(World Wide Web,简称 WWW,或 Web) 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、图象、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 是因特网的宠儿,代表着因特网信息资源的主流。它使用简单、功能强大。自 20 世纪 90 年代问世以来发展极为迅速,它的超文本、超媒体特性使之在因特网信息存储和检索领域独占鳌头。据调查,从 1993 年 6 月到 1996 年 6 月,WWW 网址的数量从 130 个增加到 23 万个,增加了 1800 倍。以下介绍几个与 WWW 相关的概念:

(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因特网上信息资源的位置。URL 包括三部分:所使用的传输协议;服务器地址;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,驻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。浏览器可分为三大类,即行式浏览器、文本浏览器(以 Lynx 为代表)和使用图形界面的浏览器。目前,广为网络用户熟悉和使用的是使用图形界面的浏览器,它们在 Windows 用户的界面下运行,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;这类浏览器有:Mosaic,Netscape,Internet Explorer (IE)等。WWW 服务器的另一优势是可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会使用 Netscape 等几个浏览器,就可满足大多数网络应用的需求。

总之,万维网信息资源是因特网信息资源的最主要、最常见的形式。

1.3.2 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、以及大型数据库等信息资源。

访问 Telnet 信息资源的一般步骤或过程是,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,有时还要回答自己所用的终端类型(如:VT100);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;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因特网其他信息服务,如:Email,FTP,Archie,Gopher,WWW,WAIS 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。

1.3.3 FTP 信息资源

FTP(file transfer protocol)是因特网使用的文件传输协议。该协议的主要功能是完成从一个系统到另一个系统完整的文件拷贝。即在因特网的联网计算机之间传输文件。前面所述的 Telnet 在远程登录后的联机状态下,用户可以浏览、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拷贝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拷贝传输到远程计算机(upload)。FTP 是获取免费软件和共享软件资源不可缺少的工具。

通过 FTP 可能获得的信息资源类型很广泛。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;美国航空航天局各种实验的信息、最新图象文件、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。

通过 FTP 获取文件的步骤一般为:

- ① 登录:旨在为 FTP 客户软件提供 FTP 服务器的地址,通过 FTP 与远程主机连接后,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,找到所需文件;
- ③ 设置文件传输参数,选择文件的传输模式。如:ASCII 模式——用于传输文本文件或二进制(Binary)模式——用于传输非文本文件。
- ④ 下载所需文件。

提供 FTP 信息资源的服务器遍布全世界,且数量不断增加,信息量日益扩大。若想有目的地查找到所需要的文件,了解该文件可从哪个匿名 FTP 服务器的哪一个目录中获得,则无异于大海捞针,必须借助相应的检索工具。因特网提供的检索工具 Archie 是专门用于搜索 FTP 服务器的地址及 FTP 文件数据库的。基于一个给定的检索项,Archie 可生成一份 FTP 文件的目录清单,包括网点的位置,准确的文件名及其所处的子目录名称。(详见第 7 章)

1.3.4 用户服务组信息资源

网上各种各样的用户通信或服务组是因特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)等等。虽名称各异,但实质上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。在这个论坛中所交流的文章即是一封封的 E-mail,因此其本质就是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,你所能想到的任何一个主题均有与之相对应的讨论组,总计已超过 15000 多个,参与者为全世界数以百万计的网络用户。

而以上述各种电子通信组形式所传递和交流的信息就构成了因特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的一个重要渠道。可以用于了解学术动态,参与讨论还可活跃自己的创新思路,