

21 世纪高等院校 电子商务教育 系列教材

电子商务 实训教程

*The Training
Teaching Material of
Electronic Commerce*

岳云康 主编

 东北财经大学出版社
Dongbei University of Finance & Economics Press

21世纪高等院校电子商务教育系列教材

电子商务 实训教程

岳云康 主编

*The Training
Teaching Material of
Electronic Commerce*

FE 东北财经大学出版社
Dongbei University of Finance & Economics Press

大连

© 岳云康 2008

图书在版编目(CIP)数据

电子商务实训教程 / 岳云康主编. —大连: 东北财经大学出版社, 2008. 3

(21世纪高等院校电子商务教育系列教材)

ISBN 978-7-81122-297-5

I. 电… II. 岳… III. 电子商务-高等学校-教材
IV. F713.36

中国版本图书馆CIP数据核字(2008)第020166号

电子商务实训教程

主编 岳云康

东北财经大学出版社出版

(大连市黑石礁尖山街217号 邮政编码 116025)

总编室: (0411) 84710523

营销部: (0411) 84710711

网址: <http://www.dufep.cn>

读者信箱: dufep@dufe.edu.cn

东北财经大学印刷厂印刷

东北财经大学出版社发行

幅面尺寸: 186mm × 230mm
2008年3月第1版

字数: 510千字 印张: 21
2008年3月第1次印刷

责任编辑: 李彬 郝树芹
封面设计: 冀贵收

责任校对: 刘东威
版式设计: 钟福建

ISBN 978-7-81122-297-5

定价: 32.00元

总序

互联网的出現為全社會提供了一種全新的商務活動方式，從而引發了對電子商務學習、實踐和培訓的熱潮。為滿足目前高等教育對電子商務教材的需求，東北財經大學出版社在 2008 年伊始開發了一套全新的“21 世紀高等院校電子商務教育系列教材”。整套系列教材圍繞電子商務的應用性知識分為三個模塊、十三種教材：第一個模塊是“原理模塊”，着力覆蓋電子商務的基本原理，包括《電子商務基礎教程》、《電子商務與網絡經濟》、《電子商務系統建設與管理》、《電子商務管理》；第二個模塊是“電子商務支持模塊”，為學習者講解對電子商務行為進行支持的主要體系，包括《電子商務案例》、《電子商務法》；第三個模塊是“電子商務中的行為模塊”，細緻刻畫了電子商務環境下的個體和企業的行為，包括《電子商務物流管理》、《電子商務交易的支付與結算》、《電子商務安全》、《電子商務網站建設與管理》、《客戶關係管理》、《網絡營銷》、《電子政務》。

這套“21 世紀高等院校電子商務教育系列教材”本着科學、先進、合理、可行的原則，在編寫過程中努力達到如下要求：

第一，博采眾長。從總體上看，由於發達國家發展市場經濟的歷史較長，市場經濟體制也比較成熟，因而其電子商務理論及相應的學科建設確實比我國領先一步，所以學習和借鑒發達國家的電子商務理論成果十分必要。同時，我國在經歷了 30 年的改革開放後，企業的體制、機制改革和技術進步已取得了巨大的成績，在電子商務實踐方面也積累了不少很有特色的成功經驗，值得總結提煉。在教材的編寫過程中，編者們廣泛參考和吸取國內外相關教材的優點，盡量做到既符合國際發展潮流，又切實反映中國電子商務實際情況。

第二，努力創新。雖然我國部分高校開辦電子商務專業的時間不長，但電子商務專業的建設已經從“摸着石頭過河”到“如何適應市場經濟中電子商務發展的需要”發生了重要的轉變。為此，電子商務專業及其教材建設在我國面臨重大變革。本套教材力求在內容和形式上都有所創新：在內容方面，更新了不適應市場經濟環境中當前電子商務實踐及未來發展的理論和觀念；在形式方面，每種教材在結構、欄目、體例及寫作風格上均有所創新，且各種教材均由“主教材”和“電子課件”兩者組成，大大方便了教與學。

第三,讲求实用。这主要表现在:一方面,内容上突出特色,兼顾理论系统性与实践可操作性。出于篇幅和知识点交叉的考虑,这套教材中每一种都力求围绕各自中心内容阐述,并根据实际课时量的要求在内容上详略得当。例如,在《电子商务基础教程》中已经详细介绍过的内容,在其他教材中就尽量避免或者简略介绍。另一方面,成熟性与创新性相结合。本次编写的教材,坚持了教材内容的成熟性与创新性的统一。在阐述成熟而稳定的教材内容的同时,适当介绍新知识、新技能、新发展趋势,使教材具有适度的超前性和前瞻性。另外,本套教材的体例要求也符合教学的规律和方法。在编写过程中,教材各章附有“学习目标”、“本章小结”、“复习思考题”、“技能实训题”等栏目,并且注重时效性。教材中的例题、案例等取材主要来自2007年全年和2008年上半年。把握本层次教材的特点,充分考虑学生吸纳知识的习惯,以本层次教学的授课时数为依据,确保教材定位准确。

第四,严把质量。本套教材由众多国内电子商务领域的专家、学者领衔编撰。他们多年从事该领域的教学与研究,具有丰富的教学及教材编写经验。他们中的大多数曾在欧美高校进修学习、合作研究或访问交流,因而对各学科的最新进展比较熟悉。他们长期关注中外企业电子商务实践,善于总结提炼。此外,各门课程教材的基本体系、结构和内容都经过各教材领衔作者的集体讨论,互提意见和建议,集思广益,严把质量关。尽管编者已经付出了最大努力,使现在所奉献给读者的这套教材体现了上述特点,但作为创新的初步尝试,难免会存在不足乃至缺陷。因此,这套教材的推出应该是任重而道远。我们希望能够尽快得到来自各方面尤其是读者方面的反馈意见,以为我们在不久的将来再版修订提供有益的参考。我们也希望并有信心通过不断修订,使教材紧随时代步伐,及时反映学科的最新进展,为培养未来的电子商务专业人才做出持续的贡献。

李强

于西安交通大学

2008年3月

前言

21 世纪是信息化的时代, 第三产业在各国的比重不断上升, 特别是服务业。信息服务业成为 21 世纪的主导产业, 这导致了电子商务的产生和发展, 在全球信息化大势所趋的影响下, 各国的电子商务不断地改进和完善, 电子商务成为各个国家和大公司争夺的焦点。

本书主要以互联网为媒介, 编排了 12 章实训内容, 每一章有任务、实训详细过程、归纳与提高、知识拓展、本章小结、复习思考题、技能实训题、小知识等栏目。12 章内容分别为: 第 1 章 网络商务信息检索与利用、第 2 章 B2C 网上商城、第 3 章 银行卡网上购物、第 4 章 第三方支付平台、第 5 章 网上拍卖、第 6 章 网上 B2B、第 7 章 CA 认证、第 8 章 电子签章和手写签名、第 9 章 BBS 与新闻组的使用、第 10 章 邮件列表的使用、第 11 章 网络广告、第 12 章 网站推广。本书的特点是实践性很强, 便于老师教学时根据需要选取有关内容。全书内容新颖、叙述严谨、实用性强, 突出了电子商务注重实践操作能力的培养和实战演练的特色。

本书适合作为电子商务、物流管理、国际贸易、市场营销、信息管理等专业的学生专业实训教材。本书从实用性出发, 力求概念准确、言简意赅、内容丰富, 不仅可以作为各大中专院校、本科院校相关专业和电子商务培训班的教材, 也可作为相关专业技术人员 and 自学者学习和参考用书。

本书由全国电子商务名师岳云康副教授担任主编, 负责拟定大纲和最后统稿、定稿。具体分工为: 山西大学商务学院岳云康老师编写前言、第 8 章和第 9 章, 山西大学商务学院白东蕊老师编写第 1、2、3、4、5、6、7 章, 中北大学分校张淑贤老师编写第 10、12 章, 山西财经大学陈亚丽老师编写第 11 章。

在本书的编写中, 作者借鉴了国内外大量的出版物和网上资料, 特对其作者表示衷心感谢, 同时, 还要特别感谢一些相关网站。

由于作者水平有限, 编写时间仓促, 书中错漏之处在所难免, 敬请广大读者批评指正。

编 者

2008 年 1 月

目录

第1章 网络商务信息检索与利用

学习目标

1.1 网络商务信息检索

1.2 网络商务信息的合理利用

1.3 归纳与提高

1.4 知识拓展

本章小结

复习思考题

技能实训题

第2章 B2C 网上商城

学习目标

2.1 任务1——实训准备

2.2 任务2——B2C 网上商城交易实例

2.3 归纳与提高

2.4 知识拓展

本章小结

复习思考题

技能实训题

第3章 银行卡网上购物

学习目标

3.1 任务1——实训准备

3.2 任务2——银行卡网上支付的准备

3.3 任务3——用银行卡网上购物

3.4 归纳与提高

3.5 知识拓展

本章小结

复习思考题

技能实训题

附录 电子商务三要素

附录A 电子商务三要素

A.1 电子商务三要素

A.2 电子商务三要素

A.3 电子商务三要素

A.4 电子商务三要素

A.5 电子商务三要素

A.6 电子商务三要素

A.7 电子商务三要素

A.8 电子商务三要素

A.9 电子商务三要素

附录B 电子商务三要素

B.1 电子商务三要素

B.2 电子商务三要素

B.3 电子商务三要素

B.4 电子商务三要素

B.5 电子商务三要素

B.6 电子商务三要素

B.7 电子商务三要素

B.8 电子商务三要素

附录C 电子商务三要素

C.1 电子商务三要素

C.2 电子商务三要素

C.3 电子商务三要素

C.4 电子商务三要素

C.5 电子商务三要素

C.6 电子商务三要素

C.7 电子商务三要素

C.8 电子商务三要素

附录D 电子商务三要素

D.1 电子商务三要素

D.2 电子商务三要素

D.3 电子商务三要素

D.4 电子商务三要素

D.5 电子商务三要素

D.6 电子商务三要素

D.7 电子商务三要素

D.8 电子商务三要素

第4章 第三方支付平台	63
学习目标	63
4.1 第三方支付平台概述	64
4.2 网银在线电子钱包	65
4.3 快钱电子商务支付平台	76
4.4 支付宝: 电子商务支付平台	82
4.5 归纳与提高	85
4.6 知识拓展	89
本章小结	93
复习思考题	93
技能实训题	93
第5章 网上拍卖	94
学习目标	94
5.1 如何卖	95
5.2 如何买	115
5.3 归纳与提高	124
5.4 知识拓展	127
本章小结	129
复习思考题	130
技能实训题	130
第6章 网上 B2B	131
学习目标	131
6.1 交易前的准备阶段	132
6.2 交易谈判和合同签订阶段	141
6.3 客户管理	150
6.4 归纳与提高	155
6.5 知识拓展	157
本章小结	163
复习思考题	163
技能实训题	164
第7章 CA 认证	165
学习目标	165
7.1 个人 CA 证书的申请及下载	166
7.2 个人 CA 证书的使用	176
7.3 企业 CA 证书	186
7.4 归纳与提高	197
7.5 知识拓展	203
本章小结	207

805	复习思考题	208
805	技能实训题	208
805	第8章 电子签章和手写签名	209
805	学习目标	209
805	8.1 任务1——电子签章的相关内容	210
805	8.2 任务2——安装电子签章、手写签名软件和制作印章	211
805	8.3 任务3——Word 签章的使用	220
805	8.4 任务4——在 Word 文档中使用手写签名	229
805	8.5 归纳与提高	230
805	8.6 知识拓展	232
805	本章小结	237
805	复习思考题	238
805	技能实训题	238
805	第9章 BBS 与新闻组的使用	239
805	学习目标	239
805	9.1 BBS 的使用	240
805	9.2 新闻组的使用	248
805	9.3 归纳与提高	257
805	9.4 知识拓展	258
805	本章小结	260
805	复习思考题	261
805	技能实训题	261
805	第10章 邮件列表的使用	262
805	学习目标	262
805	10.1 任务1——了解邮件列表	263
805	10.2 任务2——邮件列表的使用	268
805	10.3 归纳与提高	274
805	10.4 知识拓展	276
805	本章小结	278
805	复习思考题	279
805	技能实训题	279
805	第11章 网络广告	280
805	学习目标	280
805	11.1 任务1——实训准备	281
805	11.2 任务2——网络广告发布	288
805	11.3 任务3——以搜狐竞价广告为例,说明如何加入竞价广告系统	292
805	11.4 归纳与提高	293
805	11.5 知识拓展	294

298	本章小结	298
298	复习思考题	298
298	技能实训题	298
299	第12章 网站推广	299
299	学习目标	299
300	12.1 任务1——搜索引擎推广	300
307	12.2 任务2——互换链接	307
310	12.3 任务3——网络会员制营销	310
316	12.4 归纳与提高	316
317	12.5 知识拓展	317
321	本章小结	321
321	复习思考题	321
321	技能实训题	321
322	参考文献	322
323	参考网站	323

索引

第

1

章

网络商务信息检索 与利用

学习目标

- 1.1 网络商务信息检索
- 1.2 网络商务信息的合理利用
- 1.3 归纳与提高
- 1.4 知识拓展

本章小结

复习思考题

技能实训题

学习目标

知晓各种网络信息检索工具
的用法和优、劣势。

能熟练运用网络检索工具检
索商务信息。

能够合理排除干扰项，准
确、快速、全面、新颖地检索出
结果。

◆ 主要解决的问题

什么是商务信息？

有哪些网络信息检索工具？

如何准确、快速、全面、新

颖地检索出所需信息？

1.1 网络商务信息检索

Internet 是一个巨大的信息资源库，每天都有成千上万的新信息添加到 Internet。如何在众多网站中高效、快速地找到自己想要的信息，对于企业来说有着至关重要的意义。本节将介绍如何利用 Internet 快速、高效地检索网络商务信息。

1.1.1 任务1——什么是商务信息

1) 什么是信息

信息同物质、能量一样，是构成现实世界的三大要素之一，已经深入到每个人的生活、工作和学习之中。人们几乎每时每刻都在感知信息、接收信息、传播信息、利用信息。

(1) 信息的定义

由于信息涉及的分支学科比较多，所以关于信息的定义相当广泛，从不同的角度对信息可下不同的定义。目前对信息的定义已经有三十多种，分别从哲学、社会学、心理学、生物学、语言学、神经生理学、通信科学、信息科学、图书情报科学等学科角度论述了什么是信息。广义地讲，信息是物质和能量在时间、空间上定性或定量地模拟模型或其符号的集合。

信息在哲学领域中被认为是人类认识论的基础，物理学中被认为是熵的理论，数学中被认为是概率论的新发展，经济学中被认为是数据、资料、知识的统称，其中以信息论、控制论、系统论（旧三论）的角度给信息下的定义较为经典。

在控制论中，信息被理解为对外界进行调节的信号和工具，是人们为适应外界而同外界进行交换的内容和名称。维纳的定义是：“信息既不是物质，又不是能量，信息就是信息。”

信息论中，信息被认为是可以获取、变换、传递、存储、处理、识别、利用的一般对象，能为实现目标排除意外性、增加有效性。申农从通信的角度对信息下的定义是：“信息是用来减少随机不确定性的东西。”

系统论中，信息被规定为对系统的组织程度的度量。信息在系统论中概念如下：第一，信息的构成要首尾相顾，前后相连，自成一体，以便系统、全面、本质地反映客观事物，信息如不实现系统化，就谈不上系统、全面地反映客观事物。第二，信息反映要有联系性，防止孤立、静止、片面性。

(2) 信息的特点

①普遍性。信息普遍存在，并且适用于自然界、人类社会和思维以及各种学科领域。

②社会性。信息的社会性从两个方面理解：一是信息生于社会，传于社会，存于社会，用于社会，无所不在，无时不有，社会每一个成员都有权创造信息，使用信息；二是信息的内在本质是社会性的。

③共享性。信息产品在同一时间、不同地点，可以同时被大众使用。信息的共享性，可以被认为是信息有效不灭定理，或者说是信息再生定理，即信息在共享的同时，不会消失，可用数学描述为“ $1-1=1$ ”，或“ $1=\infty$ ”。

④依附性。信息必须有物质承担者，载体可以是纸张、实物、磁介质、电磁波等。

⑤消耗性。信息是有耗散和损伤的,一条信息每当被人引用或参考一次,它的价值就会减少一点,相应地耗散量就会增加。当该条信息被引用频次超过一定程度时,其使用价值也降低到一定水平之下。

⑥服务性。信息可以用于预测性、动态性、反馈性、调查研究、咨询建议、经济评价等信息服务。

⑦时效性。信息是有有效期限的,过了有效期限,信息的载体可能依然存在,但是内容已经老化。

(3) 信息、情报的区别 从认知的角度看,情报(intelligence)和信息(information)的最大区别,在于研究用户与信息系统交流是否在影响知识结构上发挥作用。情报学更专注于降低决策的不确定性,因而构成其研究的基本问题和独特性质,所有相关的技术和理论方法都围绕满足要完成的情报任务(需求的识别、表述,对象的评价、选择、学习等)而展开,要解决的是利用信息技术和理论方法实现信息的转化。而信息科学(information science)的主要任务是将信息活动的各种要素进行合理的组织和控制,从而有效地满足社会信息需求,注重的是对各要素的优化配置,从而发挥最大的整体效用。

2) 网络环境下信息资源的特点

(1) 信息的无限性

网络环境下的信息资源不仅数量庞大,而且正在以惊人的速度增长,每年都以几何数量级增长。网络环境下的信息资源因为采用结点的方式存储,因此呈现出无序性的特点。

(2) 信息的多样化

由于互联网的开发性,使得网上信息的门槛很低,各种信息层出不穷。从媒体形式上看,网络信息有文本、图像、图形、声音、视频等,即多媒体信息。从内容类型上看,网络信息有正式出版的信息,如机读数据库、电子杂志、电子报纸、电子书籍等,也有非正式出版的信息,如BBS、新闻组、专题讨论、电子邮件等。网络信息可以说是鱼龙混杂、良莠不齐,这就给想要检索到有价值信息的检索者带来一定难度。

(3) 信息机构的多样化

网络中所有能提供信息的计算机,都可看成是一个信息子网,包括专业咨询机构网站、ICP、书目信息中心(如OCLC)、文献出版机构、学术科研机构和其他提供信息的相关机构等。如何选择合适的信息机构进行访问,对于信息检索者来说,是一个首先需要思考的问题。

(4) 信息的时效性更强

互联网上的信息几乎每时每刻都在更新和淘汰,我们面对的是一个日新月异的信息世界。我们也可以在网看到最新的第一手资料。

3) 网络商务信息

对于现代企业来说,信息可看作是企业的生命,是企业不可须臾离开的法宝。网络商务信息,不仅是企业进行网络营销决策和计划的基础,而且对于企业的战略管理、市场研究以及新产品都有着极为重要的作用。

相对于传统商务信息,网络商务信息具有以下显著的特点:

(1) 实效性强

传统的商务信息,由于传递速度慢、传递渠道不畅,经常导致“信息获得了但也失效了”。

网络商务信息则可有效地避免这种情况。由于网络信息更新及时、传递速度快,只要信息收集者及时发现信息,就可以保证信息的实效性。

(2) 准确性高

网络商务信息的收集,绝大部分是通过搜索引擎找到信息发布源获得的。在这个过程中,减少了信息传递的中间环节,从而减少了信息的误传和更改,有效地保证了信息的准确性。

(3) 便于存储

现代经济生活的信息量是非常大的,如果仍然使用传统的信息载体,把它们都存储起来难度相当大,而且不易查找。网络商务信息不仅可以方便地从互联网上下载到自己的计算机上,通过计算机进行信息管理,而且在原有的各个网站上,也有相应的信息存储系统,自己的信息资料丢失后,还可以到原有信息源中再次查找。

1.1.2 任务2——如何通过网络检索商务信息

1) 在网上检索所需要的信息

(1) 浏览专业网站的信息,如中国经济信息网(<http://www.cei.gov.cn>)、中国广告信息网(<http://www.chinaad114.com>)、中国商品交易市场(<http://www.chinamarket.com.cn>)、中国机械工商网(<http://www.chinajx.com.cn>)。对于专业检索人员,脑海里应随时储备有不同领域的门户网站,并不断加以更新和完善。

(2) 通过专业数据库,如联机检索数据库 dialog、万方数据(<http://www.wanfang.com.cn>)、中国知网(<http://www.cnki.com.cn>)。

(3) 利用 E-mail 或站点法进行调研。网上调查问卷可以直接根据传统的市场调查问卷形式制作。调查问卷可由多个问题组成,问题可包括需用户输入信息的填空问题、单项选择问题、多项选择问题,并可指定必答项和非必答项。调查问卷生成前和生成后都可即时修改。调查问卷应清楚写明自己企业的所在地、通讯地址和联系方式。

(4) 在相应网站的 BBS 上与网友交流。

(5) 利用域名公式“www. 公司站点类型”,猜测待查的组织的站点,直接到其站点查找有关信息。例如,若想知道 Dell 公司的产品信息,可直接到站点 <http://www.dell.com> 查找;若想知道有关人口统计的信息,可直接到站点 <http://www.census.gov> 查找。

(6) 使用搜索引擎。互联网的发展大大促进了信息检索技术的发展和运用,一大批搜索引擎产品的诞生,为网民提供了信息导航和快速获取的工具。

虽然网络系统提供了许多检索方法,但堆积如山的、全球范围的、各行各业的信息,常常把企业营销人员淹没在信息的海洋之中。在浩瀚的网络信息资源中,迅速地找到自己所需要的信息,经过加工、筛选和整理,把反映商务活动本质的、有用的、适合本企业情况的信息提炼出来,需要相当一段时间的培训和经验积累。

2) 网络信息检索的步骤

(1) 听取用户要求;

(2) 全面调查企业现状、环境、竞争对手、产品状况等;

(3) 分析企业现实信息需求和潜在信息需求;

(4) 利用网络检索工具查找到调研信息;

(5) 对收集到的信息进行筛选、归类、汇总、整理;

信息(6) 反馈给用户, 交流意见; 对检索结果进行评价, 信息的中点故 WWW 是其重要(7) 调整检索策略和调查方式, 在原有基础上进行更深层次的挖掘。

3) 网络信息检索的特点

(1) 交互式作业方式。所有的网络信息检索工具都具有交互式作业的特点, 因此具有良好的信息反馈功能和瞬间反应功能。信息反馈和瞬间反应是传输信息检索系统性能的最重要指标, 在网络环境下也具有同样的意义。

(2) 用户透明度。网络信息检索对用户屏蔽了网络的各种物理差异, 使用户在使用这些服务时感受到明显的系统透明度。这里所指的物理差异包括主机的硬件平台、操作系统等软件上的差异、客户程序和服务程序版本上的差异、主机的地理位置的差异、信息的存储方式甚至通信协议的差别(如 WWW 客户程序可以通过多种协议使用各种不同的信息资源)等。这一特点对网络环境下的信息检索来说是十分关键的。

(3) 信息检索空间的拓宽。信息检索空间是衡量信息检索工具的重要指标之一。网络信息检索在这方面具有传统信息检索和 Internet 基本信息服务所不具备的优势。以 FTP 为例, 尽管使用 FTP 可以检索所有的 FTP 服务器, 但是用户必须预先知道这些服务器所在的主机地址, 而且在某一时刻只能使用一个 FTP 服务器。网络信息检索工具的工作方式则与此不同, 它们可以同时使用多个主机甚至是所有主机的某种资源, 而且用户不必知道它们的具体地址。这一特点为用户带来的好处是显而易见的。

(4) 友好的用户界面。与 Internet 的三大基本信息服务相比, 网络信息检索系统的用户界面要友好得多, 特别是一些商业化软件(如 Internet Explorer 和 Netscape Navigator)。即使是 Internet 上的一些免费软件(如 MS Windows 和 Unix 下的各种服务程序和客户程序)也设计得相当不错。对于有一定微机使用经验的人来说, 学会使用这些软件是轻而易举的事情。Internet 的普及在很大程度上是得益于这些设计精良的软件。

小常识 1-1

信息检索的含义

信息检索 (information retrieval), 通常指文本信息检索, 包括信息的存储、组织、表现、查询、存取等各个方面, 其核心为文本信息的索引和检索。从历史地看, 信息检索经历了手工检索、计算机检索到目前网络化、智能化检索等多个发展阶段。

信息检索的对象从相对封闭、稳定一致、由独立数据库集中管理的信息内容扩展到开放、动态、更新快、分布广泛、管理松散的 Web 内容; 信息检索的用户也由原来的情报专业人员扩展到包括商务人员、管理人员、教师、学生、各专业人士等在内的普通大众, 他们对信息检索从结果到方式提出了更高、更多样化的要求。适应网络化、智能化以及个性化的需要是目前信息检索技术发展的新趋势。

1.1.3 任务3——学会使用搜索引擎

下面我们最受网民欢迎的搜索引擎来介绍网络商务信息资源的检索。

搜索引擎采用的是全文检索技术, 是互联网提供公共信息检索服务的 Web 站点。它是具有一定的技术和策略在互联网中搜索, 发现网络信息, 并对网络信息进行理解、提取和处理, 为网络用户提供检索服务, 从而快速查找检索信息的一种网络工具。其主要任务是在 Internet 上主动搜

索其他 Web 站点中的信息,并对其自动索引,其索引的内容存储在一个大型数据库中。当用户利用关键词查询时,该网站会返回包含该关键词的所有已收录在数据库中的网址信息,并提供通向该网站的链接。

1) 搜索引擎的类型

世界上三大著名的商业搜索引擎公司是 Yahoo、Excite 和 Altavista。一般来说,搜索引擎根据检索形式、对象和范围等的不同,可以分为多种不同的类型。情报检索界通常是按索引数据库中数据的搜集和加工组织方式的不同,将其大体分为分类目录式搜索引擎和关键词式搜索引擎两类。

(1) 分类目录式搜索引擎

分类式搜索引擎主要采用人工方式来建立和维护其分类目录,按照某些分类形式组织信息资源。目录按照一定的主题分类体系组织,并辅之年代、地区等分类。用户一般采取逐层浏览目录、逐步细化来寻找合适的类别直至具体资源。分类式搜索引擎的优点在于系统、准确、结构清晰明了、导航质量高。分类式搜索引擎对于关键词不太确定的资料查询十分有效。如果想找到一些相关性较大的信息群体,或某一方面比较系统的信息资源,就可以使用目录式搜索引擎,一级一级地查找。如 Yahoo 就是结合传统分类法和网络信息的特点,把网上的各种资料分得很细,有休闲运动、健康医药、艺术人文等很多类别;其不足之处在于制作起来投入人力较多,搜集、组织和维护信息较为困难,对网页信息更新的实时反应能力较差。这类搜索引擎的代表有雅虎、搜狐等。

(2) 关键词式搜索引擎

关键词式搜索引擎是搜索引擎和网络检索技术的核心,主要采用自动搜索与标引方式来建立和维护其索引数据库。用户查询时可以用逻辑组合方式输入关键词,而且可以使用逻辑算符、位置算符、截词符等来构成检索式,可以限制检索对象的地区、数据类型、时间等,因而可准确检索满足特定条件的网络资源。关键词式搜索引擎通过特定的检索软件,查找其索引数据库,给出与检索条件相匹配的检索结果,供用户浏览使用。这类搜索引擎的优点在于制作时不需要很多的人工介入,搜索范围广,信息量大,更新及时。用户使用时,简单、明了、直观。这类搜索引擎特别适用于用户不知道其要查找的内容是属于哪个类目的情况,或者要搜索的信息比较冷门,没有被搜索引擎收录的情况。其缺点在于返回信息较多,有很多信息噪音,查准率较低,用户有时难以用几个关键词表达清楚自己的意思。这类搜索引擎较为著名者有百度、天网、悠游等。

目前,分类目录式搜索引擎与关键词式搜索引擎正在向融合的趋势发展,二者互相结合、互相补充、取长补短,从而更好地为网络用户服务。很多门户网站或搜索引擎都采用了两种方式相结合的手段。

按照搜索引擎采用的语言分类,搜索引擎主要有英文搜索引擎和中文搜索引擎两类。常用的英文搜索引擎有 Google、Yahoo、Infoseek、Altavista 等。常用的中文搜索引擎有中文 Google、中文 Yahoo、Baidu、Sogou、Sohu 等。用户应能掌握几种常用的搜索引擎的用法。

2) 几种中文搜索引擎的比较

新浪是全球最大的华人网站,其数据库中收录了 200 多万个网页。新浪搜索引擎提供关键词查询和分类检索两种查询方法,信息采集方式为网络机器人自动搜索,是目前检索软件中功能最全面、查全率最高的优秀搜索引擎之一。

搜狐的搜索引擎是以提供分类目录为主的中文搜索引擎,其分类原则是以图书分类为基础,

与日常应用习惯相结合,由编辑人员进行分类。搜狐作为一种搜索引擎,其系统反应速度快,查询准确率高,便于简单查询。

网易的搜索引擎很有特色,它先将用户的检索式在自己的分类库中进行查询,如果没有检索出结果,系统自动转向全文数据库进行搜索。如果在分类库中检索出结果,用户对检索结果不满意,可以直接按检索结果页面底部的全文检索按钮,继续在全文库中进行检索。这对于一个非专业用户来说,是非常实用的一种检索策略。

FM365 号称是新三大门户网站之一,提供分类目录式检索和关键词式检索两种途径,其搜索能力强,反馈信息也很丰富。查询结果质量相关度很高,且支持模糊搜索。

Yahoo 中文搜索引擎收录了全球咨询网上数以万计的中文网址,提供按目录逐级搜索和输入关键词检索两种检索方式,信息采集方式由索引人员用人工方式建立并更新。它支持简单和高级查询功能,其系统反应速度很快,界面友好,而且通过主题指南查询准确率高,内容丰富。另外,针对目前中文的状况,雅虎分别设立了简体中文(GB)和繁体中文(BIG 5)两种版本。

天网搜索引擎,是国家“九五”科技攻关重点项目“中文编码结合分布式中英文信息发现”的研究成果。它采用“汉语词切分与词性标准软件”的技术,准确率达到 95%,能自动收集、识别互联网上的 WWW、Newsgroup 和 FTP 等多种格式信息,智能化地提取摘要和关键词,建立索引,提供查询和对不良信息进行监控与报警功能的网络信息自动发现和跟踪服务。

Google 搜索引擎是收集亚洲网站最多的搜索引擎之一,也是目前最好用的搜索引擎之一。它提供了一系列革命性的新技术,包括完善的文本对应技术和先进的 PageRank 排序技术。不同于大多数搜索引擎将链接文本和它所在的页面相关联的方式,Google 搜索引擎把链接文本和它指向的文档联系在一起,从而进一步改善了查询质量。但美中不足的是,其数据的更新速度无法进一步提高,数据量的庞大使得 Google 搜索引擎的数据更新无法早于 30 天,这在一定程度上影响了用户对信息的时效需求。

1.1.4 任务4——关键词的选择

1) 选择适当的查询词
选择适当的查询词是靠经验积累,当然我们也可以摸索一些规律。

(1) 查询词表述准确

查询词表述准确是获得良好搜索结果的必要前提。例如,要查找 2005 年国内十大新闻,查询词可以是“2005 年国内十大新闻”,但如果把查询词换成“2005 年国内十大事件”,搜索结果就没有能满足要求的了。

另外,对于一些有歧义的词,特别要注意表述出我们需要的意思。例如,要查找绿色食品,这个“绿色”就是一个非常容易引起歧义的词语,一方面它可以表示绿颜色的食物,如绿色蔬菜,另一方面它也可以表示无污染的健康食品。这时候就需要在输入时把它的内涵意思诠释出来作为查询词输入。

(2) 查询词的主题关联与简练

目前的搜索引擎并不能很好地处理自然语言。因此,在提交搜索请求时,最好把自己的想法提炼成简单的,而且与希望找到的信息主题相关联的查询词。

例如,某三年级小学生,想查找一些关于时间的名人名言,他的查询词是“小学三年级关于时间的名人名言”。这个查询词很完整地体现了搜索者的搜索意图,但效果并不好。“小学三