



全国高校素质教育教材研究编审委员会审定

21 世纪创新型人才培养高职高专“信息技术”系列规划教材

网络信息检索技术 实训教程

谭力红 主 编



科学普及出版社

全国高校素质教育教材研究编审委员会审定

21 世纪创新型人才培养高职高专“信息技术”系列规划教材

网络信息检索技术 实训教程

谭力红 主 编

林向建 麦 影 副主编

科学普及出版社

• 北 京 •

图书在版编目 (CIP) 数据

网络信息检索技术实训教程/谭力红 主编. —北京: 科学普及出版社, 2007.7

ISBN 978-7-110-06611-9

I. 网… II. 谭… III. 计算机网络—情报检索—教材

IV. G252.7

中国版本图书馆CIP数据核字 (2007) 第088635号

自2006年4月起本社图书封面均贴有防伪标志, 未贴防伪标志的为盗版图书。

网络信息检索技术实训教程

谭力红 主编

责任编辑: 周晓慧 高立波

责任印制: 王 沛

封面设计: 张骐年

出版发行: 科学普及出版社

社 址: 北京市海淀区中关村南大街 16 号 邮编: 100081

电 话: 010-62103210 传真: 010-62183872

排 版: 科士洁文印中心

印 刷: 北京正道印刷厂

开 本: 787mm × 1092mm 1/16

印 张: 12.5

字 数: 290 千字

版 次: 2007 年 7 月第 1 版

印 次: 2007 年 7 月第 1 次印刷

印 数: 1-2500 册

书 号: ISBN 978-7-110-06611-9/TP · 177

定 价: 25.80 元

版权所有 翻印必究

如有印装质量问题, 请将本书寄回编委会由我们负责为您调换

地址: 北京市海淀区交大东路 62 号西楼 208 室 100044

内 容 提 要

本书包含了数字化图书、期刊资源的检索，专利、标准、科技成果的检索，政策法规、企业与机构信息以及商务信息等方面的检索等内容。

全书分为两部分：第一部分的检索教程共五章，第一章介绍了计算机信息检索的基本知识；第二章至第五章使用了丰富的案例图片，详尽地叙述了目前广泛使用的中文数据库和中文搜索引擎及其操作流程以及网络信息检索实训的操作步骤，力求信息检索过程的实用性和可操作性。第二部分的检索实训强化案例教学，使学习者进一步实践信息检索的过程，体现信息检索课程的实践性和开放性。

本书可以作为高职高专相应课程的教材，计算机信息技术培训教材以及网络爱好者的参考资料；也可以作为高校学生准备毕业论文写作，或科研人员研究论文写作时的信息检索指南；同时适合从事公务、商务和科研等社会活动的人们能及时有效地获取与自身需求相关的信息。

（如果需要课件，请 E-mail: lxj88@Tom.com）

序

21 世纪,人类社会已进入到信息化时代。在信息社会中,物质世界正在隐退到信息世界的背后,人们进入了一个信息的海洋,各类信息组成了人类的基本生存环境。现在,《纽约时报》一天的信息量等于 17 世纪一个人一生所能得到的信息的总和;因特网平均每隔 30 秒钟就有一台计算机加入。面对爆炸式增长的信息量,人们犹如逆水行舟,不及时有效地处理和吸取自己所需的信息,就将成为这个时代的弃儿。另一方面,信息又是现代社会的宝贵资源,随着社会的发展进步,信息的价值将不断飚升。正如世界著名未来学家阿尔温·托夫勒所预言的:“谁掌握了信息,控制了网络,谁就掌握了整个世界。”

早在信息洪流到来之时的 1974 年,美国信息产业协会主席保罗·泽考斯基就提出了“信息素养”(information literacy)这一重要命题。尽管人们对“信息素养”的内涵理解各异,信息素养在不同层次的人身上体现的侧重面也不一样,但概括起来,它主要包含两个层面和具有五大特征。两个层面即人文层面和技术层面:从人文层面来看,信息素养反映了人们面对信息的心理状态,也可以说是面对信息的修养,包括有获取新信息的意愿,能够充满自信地运用各类信息解决问题,有较强的创新意识和进取精神等。从技术层面来讲,信息素养反映的是人们利用信息的意识和能力;信息素养的五大特征:即捕捉信息的敏锐性、筛选信息的果断性、评估信息的准确性、交流信息的自如性和应用信息的独创性。

教育要对 21 世纪信息社会的到来做出积极反映。由全国高校素质教育教材研究编审委员会组织研编的《21 世纪全国高校创新型人才培养规划教材——“信息技术”系列》,以创新型人才培养为目标,以掌握信息技术为导向,顺应时代、高瞻远瞩地回应信息时代的教育诉求。信息技术是对信息的获取、理解、分析、加工、处理、传递等有关技术的总称,其核心问题是信息能力的培养。所谓信息能力就是对各种信息技术的理解和活用能力,即对信息的获取、理解、分析、加工、处理、创造、传递的理解和活用能力,它是信息时代的一种必备能力。而信息能力又是一种基本的信息素养,它是一种在技术层面、操作层面和能力层面上的素养,它是信息素养的重要内容。所以,培养学生的信息能力、提高学生的信息素养是信息技术教育的根本目标。概言之,我们所理解的信息技术教育,应该是一种以培养学生信息能力,提高学生的信息素养为目标的学科教育,它不但是人们生存于信息时代的当务之急,更是实现终身学习的必经之途。

《网络信息检索技术实训教程》,作为上述《21 世纪全国高校创新型人才培养规划教材——“信息技术”系列》教材之一种,主要讲述“信息技术”中的一种重要技术——“信息检索技术”,而且是基于现代网络的信息检索技术。说“信息检索技术”是“信息技术”的重要环节,是因为我们面对信息爆炸时代的降临,面对海量信息的重重包围,决定人生成败与否的关键因素首先是对信息的“检索”与处理。换言之,培养学生在信

息时代生存的基本技能,就是应该促使学生掌握“搜寻”自己所需的信息并能对已有信息进行有效地组织、整理、利用和传播的能力和方法。所以说“信息检索技术”是一种重要的“信息技术”,“信息检索能力”是一种重要的“信息能力”,它们都属于基本的“信息素养”。

基于此,我国有学者最近提出了一个新的概念——搜商。所谓“搜商”,顾名思义,是人们借助工具获取新知识的能力,即搜索能力。具体地说,是指人类使用搜索引擎在互联网的浩瀚数据海洋中寻找所需信息的能力。它与智商、情商相并列,构成人类重要的智力因素。100年前,法国心理学家比奈提出“智商(IQ)”;10年前,哈佛大学心理学教授丹尼尔·戈尔曼提出“情商(EQ)”;今天,“搜商(SQ)”概念的提出也不是偶然的,是人类信息技术发展的必然产物。从问路到查地图,从传统的“手工文献检索”到现代的“网络信息检索”,无一不是人们随着科学技术的进步对知识、信息的选择、搜寻过程,亦是人类“搜商”的发展和提高过程。只是到了信息爆炸的今天,随着数字科技的发展特别是互联网和搜索引擎的诞生,给了人们一个至今为止获取知识最为快捷、最为强大的工具。这同样也使人们获取知识的能力在现代科技的放大作用下产生了显著的差异。所以,搜商和一个人运用搜索引擎的熟练程度、对不同搜索引擎特点的了解程度以及自身掌握知识面的广度都有莫大的关系。在这个意义上,“网络信息检索技术”的学习可以说旨在提高学习者的“搜商”和“信息素养”,在信息社会,掌握获取知识的能力比获取知识更为重要。

《网络信息检索技术实训教程》即为上述目的所编。主编谭力红副教授从事《文献检索》教学与科研经过了跨越世纪的十余年,本书可以说概括了她十余年的教学成果和科研心得。《网络信息检索技术实训教程》遵循应用性、实践性和技能性的原则,既体现了教材的科学性、逻辑性和系统性,又强化了教学内容与信息技术的整合性,教学过程与教学环节的开放性以及教学形式与教学手段的实训性等特点。不仅适用于高职高专学生学习,也可以作为计算机信息技术培训的教材。

颜泽贤

2007年春于羊城

前言

21 世纪的今天,人类进入了一个崭新的网络信息时代,因特网成为世界上最大的信息资源库。编者从 1992 年起讲授《文献检索》课程至今,跨过了世纪的时空,经历了《文献检索》课程从手工检索到计算机网络检索的每一进步。今天的网络信息检索代表着在网络信息环境下发展起来的一种新型检索模式,已经成为信息检索领域的重要分支。今天的信息检索技术应该成为每一个人所具备的基本技能。《网络信息检索技术实训教程》编写的目的,是让学习者学会在浩瀚的信息海洋中及时有效地获取与自身需求相关而且适用的信息,并在现代社会的经济活动和科技活动中提高个人的竞争力。

本书遵循“应用性、实践性和技能性”的原则,在原有信息检索教材的科学性、逻辑性和系统性的基础上,具有以下特点:

1. 强化以问题为中心的案例教学

尽量减少大篇幅地论述理论知识,以检索实例为纲,在案例中串联基本理论知识,用一个或几个案例解决一个问题,使学习者在阅读、实践案例的过程中得到技能的培养与提高。

2. 强化教学过程的实践性和开放性

结合不同的专业,精心设计检索实训题目。使学习者以课题为纲,把信息检索技术的知识、技能融会贯通到解决问题的过程之中。答案是开放的,如检索思路的表达,检索方案的设计,检索工具的选择,检索结果的归类,信息内容的评价与处理等。

3. 强化教学内容与信息技术的整合

教学的重点部分:信息检索部分在网上进行,使学习者学会在信息技术条件下有效地获取所需的信息,并融入自己的知识基础,能有效地利用信息去完成一个具体的任务。

构建教学内容与信息技术整合的新型态,使学习者成为信息技术最大程度的应用者和受益者,这是中国高等教育学会“十一五”立项课题“信息技术环境下具有高职特色的整合型课程新型态”(编号:06AIQ0130048)的核心思想,本书是该课题的研究成果之一,在编写上体现一种新的教学模式。

本书第一部分的第一、第二、第四章由谭力红编写;林向建编写第三章,并统编全书;第五章由麦影编写。第二部分的检索实训由谭力红、林向建和麦影共同设计完成。

感谢广东省社会科学联合会主席颜泽贤教授为本书作序,他精辟的论点“搜商——信息社会的信息素养”深刻地告诉人们:“谁掌握了信息,控制了网络,谁就掌握了整个世界。”

感谢我校图书馆潘妙辉馆长对本书编写的指导和帮助。

感谢我校成建姬、章玉平、周会青等老师对本书编写提出指导性建议。

感谢彭强主任、潘小珍主任对本书编写工作的大立支持。

谭力红

2007 年春于广州城市职业学院

目 录

第一章 计算机信息检索概论.....	1
1.1 文献与信息检索.....	1
1.1.1 文献	1
1.1.2 信息检索的涵义.....	1
1.1.3 信息检索方式、检索工具及其类型.....	2
1.1.4 信息检索的类型.....	3
1.1.5 信息检索语言.....	3
1.1.6 文献检索的常用途径.....	5
1.1.7 文献检索的方法.....	5
1.1.8 查全率与查准率.....	6
思考与练习	7
1.2 计算机信息检索基础.....	8
1.2.1 计算机信息检索的含义及特点.....	8
1.2.2 计算机信息检索的原理：存储与检索	8
1.2.3 计算机信息检索系统的构成.....	8
1.2.4 计算机信息检索的基本技术.....	9
1.2.5 计算机检索的技巧和方法.....	11
思考与练习	11
第二章 中文数据库	13
2.1 万方数据资源系统.....	13
2.1.1 万方数据资源系统子系统.....	13
2.1.2 快速检索步骤.....	14
2.1.3 跨库检索	17
思考与练习	25
2.2 中国学术期刊数据库（CNKI）	32
2.2.1 初级检索	32
2.2.2 高级检索	40
2.2.3 专业检索	43
思考与练习	46
2.3 维普资讯网	49
2.3.1 文章搜索	49
2.3.2 分类检索	55

2.3.3 期刊导航	57
2.3.4 《中文科技经济新闻数据库》	59
思考与练习	63
2.4 超星电子图书	65
第三章 网络信息资源及其检索	72
3.1 网络信息资源及其组织方式	72
3.1.1 网络信息资源的定义	72
3.1.2 网络信息资源的特点	72
3.1.3 网络信息资源及检索方法	73
3.2 网络信息资源检索工具——搜索引擎	74
3.2.1 搜索引擎	74
3.2.2 搜索引擎的一般查询规则	75
3.3 中文搜索引擎	85
3.3.1 综合型检索工具 Google	85
3.3.2 综合搜索引擎 雅虎 (Yahoo)	90
3.3.3 百度搜索引擎	92
3.3.4 搜狐搜索引擎	93
3.3.5 天网搜索引擎	94
3.3.6 赛迪网“IT 罗盘”搜索引擎	94
3.3.7 万纬搜索引擎	98
第四章 科技信息资源检索	103
4.1 国家科技图书文献中心	103
4.1.1 快速检索步骤	104
4.1.2 数据库检索	106
4.2 中国科技信息网	117
4.3 中国技术创新信息网	123
4.4 专利文献检索	128
4.4.1 中国专利信息网	128
4.4.2 中华人民共和国国家知识产权局	136
4.5 标准文献检索	141
4.5.1 中国标准服务网	141
4.5.2 中国标准化信息网	144
4.5.3 中国质量信息服务网	145
4.5.4 国际标准化组织	145
4.5.5 标准全文的获取	146
思考与练习	146

第五章 经济信息资源检索	147
5.1 中国经济信息网.....	147
5.1.1 概述	147
5.1.2 综合篇	148
5.1.3 企业篇	151
5.2 国务院发展研究中心信息网.....	155
5.2.1 概述	155
5.2.2 全文检索	155
5.2.3 金融版	157
5.2.4 企业版	158
5.3 中国宏观经济信息网.....	160
5.3.1 概述	160
5.3.2 信息检索	160
5.3.3 “中宏速递”栏目.....	161
5.3.4 “形势分析”栏目.....	162
5.3.5 “中宏特稿”栏目.....	163
5.4 阿里巴巴(中国)	164
5.4.1 概述	164
5.4.2 网上企业的建立及运作.....	164
5.4.3 商业信息检索.....	168
5.4.4 网上订购	170
5.4.5 商业信息的分类浏览及检索.....	171
网络信息检索实训 1: 万方数据库的使用	174
网络信息检索实训 2: CNKI 数据库的使用	176
网络信息检索实训 3: 维普资讯网的使用	178
网络信息检索实训 4: Google 搜索引擎	180
网络信息检索实训 5: 课题检索	182
网络信息检索实训 6: 网络资源搜索	183
网络信息检索实训 7: 主题检索	184
网络信息检索实训 8: 搜集某产品的市场信息	185
参考文献	186

第一章 计算机信息检索概论

1.1 文献与信息检索

1.1.1 文献

1. 定义

“文献”一词在1983年颁布的国家标准《文献著录总则》中将其定义为：记录有知识的一切载体。

文献的定义包含文献的三要素：

- (1) 知识信息内容，是文献的灵魂所在；
- (2) 载体材料，即可以提供记录知识或信息的物质材料；
- (3) 记录方式，即用文字、图形、符号、声音和视频等方式和技术手段把知识或信息记录在一定的物质载体上。

2. 分类

文献按其加工程度分类为：

- (1) 一次文献或原始文献。例如专著、报刊、论文、会议文献、学位论文等；
- (2) 二次文献，又称检索性文献。是在一次文献的基础上加工而成的文献，具有揭示文献内容特征（主题、分类）或外表特征（作者、篇名）的功能。例如目录、索引、文摘、题录等；
- (3) 三次文献，指利用二次文献的线索，系统地检索出一批相关文献，并对其内容进行综合、分析、研究和评述而编写出来的文献。例如综述、工具书等。

1.1.2 信息检索的涵义

广义的信息检索：是指将信息按一定的方式组织和存储起来，并根据用户的需要找出有关信息的过程，所以它的全称又叫“信息的存储与检索”。

狭义的信息检索则仅指该过程的后半部分，即从信息集合中找出所需要的信息的过程，相当于人们通常所说的信息查询。

完整的信息检索概念包括两项工作：信息存储和信息检索。

信息存储过程：

即搜集、筛选、整理、积累现有信息及检索工具，建立信息检索系统。

信息检索过程：

即利用信息检索工具查找所需要的信息。

1.1.3 信息检索方式、检查工具及其类型

1. 检索方式

(1) 手工检索：人们直接查找印刷型文献的检索。

(2) 计算机检索：通过计算机对已经数字化的信息，按设计好的程序进行查找和输出的过程。

2. 检索工具

信息检索工具是指按某种方式、方法建立起来的供读者查检信息的一种有层次的体系，是表征有序的信息特征的集合体。其中二次或三次文献是文献信息系统的核心和概括。

其功能为报道文献信息、存储文献信息、检索文献信息。

3. 信息检索工具类型

(1) 按照信息存储和检索所采用的设备和手段划分：手工信息检索系统：书本式检索系统；卡片式检索系统。

计算机信息检索系统：光盘检索系统；联机检索系统；网络检索系统。

(2) 按揭示信息内容的程度划分：

1) 目录：目录是以一件或一种完整的出版物（如一本书、一种期刊等）作为著录基本单位的检索工具。一般目录主要揭示、报道这些单位出版物的外表特征。例如：

表 1-1-1 目录格式

《广州城市职业学院学报》2006 年第 2 期目录（局部）				
风险社会				
	科学技术社会风险的价值反思		/李建珊 乔文娟	(1)
经济管理				
	中国如何在第二次全球化过程中打造新经济版图		/王 廉	(22)
教育教学				
	我国高校本科专业设置改革探讨		/刘楚佳	(52)
	高职实践教学教材的特点与编制原则		/吴志柳 罗成龙	(57)

2) 题录：题录是以单篇或单份文献为著录的基本单位的检索工具，题录也主要揭示、报道文献的外表特征，但其著录格式与目录有所不同。例如表1-1-2：

表 1-1-2 题录格式

序 号	篇 名	作 者	刊 名	年/期
1	浅论数据库的建设与应用	乜 勇	中国医学教育技术	2006/05
2	医院 PACS 系统的效益研究	端 妮	中国医学教育技术	2006/05
3	给你的课件配个快速导航菜单——Authorware 中右键快速导航菜单程序的设计与制作	田保红	中国医学教育技术	2006/05
4	基于 ASP 技术的网络考试组卷和评分系统的设计与实现	杨 蕾	中国医学教育技术	2006/05
5	运动解剖学多媒体课程建设	柏友萍	中国医学教育技术	2006/05

3) 索引: 索引是根据一定的需要, 把特定范围内的某些重要文献中的有关款目或知识单元(如书名、刊名、人名、地名、语词等), 按照一定的方法编排, 并指明出处, 为用户提供文献线索的一种检索工具。

表 1-1-3 索 引

序 号	刊 名	ISSN
1	传感器与微系统	1000-9787
2	传感技术学报	1004-1699
3	传感器技术	1000-9787
4	工矿自动化	1671-251X
5	环境遥感	

4) 文摘: 文摘是在题录的基础上, 在每条著录款目后边再加上文献内容的摘要。

表 1-1-4 文摘格式

计算机软件类课程教学模式探讨

【作者】	徐用高;
【作者单位】	长江大学计算机科学学院; 湖北荆州;
【刊名】	中国医学教育技术, <i>China Medical Education Technology</i> , 2006 年 05 期 期刊荣誉: ASPT 来源刊 CJFD 收录刊
【关键词】	软件类课程; 三段式教学模式; 作品分析法; 课件;
【摘要】	简要介绍了计算机软件类课程的特点, 提出了计算机软件类课程“三段式”教学模式, 并详细论述了该模式的实施步骤和方法, 最后指出该模式在应用中存在的问题和应对策略。

1.1.4 信息检索的类型

1. 文献检索

例如, 查找“计算机网络安全”方面的研究信息。

2. 数据检索

例如, 了解2005年广州GDP数字, 某电压调节器的参数。

3. 事实检索

例如, 了解有哪些机构在进行计算机网络安全的研究。

4. 图像、声音检索

例如, 查找旧广州西关大屋的照片, 民国时期广州长堤建筑物的图片, 毛泽东的讲话录音。

1.1.5 信息检索语言

1. 定义

检索语言是指根据信息检索的需要而创造的专供信息存储和信息检索使用的规范化

的一种人工语言，它贯穿于信息存储和检索的全过程。

信息检索语言是沟通信息存储和信息检索的一种约定语言。

2. 检索语言的种类

表述文献外表特征的语言：题名，责任者，号码，引文。

表述文献内容特征的语言：

(1) 分类语言：按知识门类的逻辑次序将信息系统地组织和划分的语言，具体体现为用分类号和类名来表达信息的主题概念。如《中国图书馆图书分类法》、《国际十进分类法》、《美国国会图书馆图书分类法》等。

《中国图书馆图书分类法》（第4版）采用五分法，即分5个基本部类，各部类下再展开为22个基本大类，各基本大类下再分二级、三级、四级、五级子类目。

1) 基本部类有五类：马克思主义、列宁主义、毛泽东思想；哲学；社会科学；自然科学；综合性图书。

2) 基本部类下分为 22 个大类，它们的标识符和类名如表 1-1-5 所示。

表 1-1-5 22 个大类的标识符和类名

A. 马克思主义、列宁主义、毛泽东思想	N. 自然科学总论
B. 哲学	O. 数理科学和化学
C. 社会科学总论	P. 天文学、地球科学
D. 政治、法律	Q. 生物科学
E. 军事	R. 医药、卫生
F. 经济	S. 农业科学
G. 文化、科学、教育、体育	T. 工业技术
H. 语言、文字	U. 交通运输
I. 文学	V. 航空、航天
J. 艺术	X. 环境科学
K. 历史、地理	Z. 综合性图书

3) 22 个大类下细分构成简表，也称二级子类目（如表 1-1-6 所示），简表细分又构成详表，又称三级子类目。

表 1-1-6 T 大类的二级子类目

T	工业技术	TL	原子能技术
TB	一般工业技术	TM	电工技术
TD	矿业工程	TN	无线电电子学、电信技术
TE	石油、天然气工业	TP	自动化技术、计算机技术
TF	冶金工业	TQ	化学工业
TG	金属学与金属工艺	TS	轻工业、手工业
TH	机械、仪表工业	TU	建筑科学
TJ	武器工业	TV	水利工程
TK	能源与动力工程		

(2) 主题语言: 是一种描述性语言, 还可细分为标题词语言、关键词语言和叙词语言。

标题词语言——是从文献的题目和内容中抽出来, 经过规范化处理的主题语言。

关键词语言——直接从文献信息的标题、摘要或内容本身抽取出来的用于揭示信息主题内容的自由词。

叙词语言——是以表达文献主题内容的概念单元为基础, 经过规范化处理, 可以进行逻辑组配的一种主题语言。

(3) 名称语言: 人名、地名、书刊名等。

(4) 代号语言: 专利号、标准号、合同号等。

1.1.6 文献检索的常用途径

1. 作者途径

通常按姓在前名在后的方式检索。

2. 分类途径

我国高校与公共图书馆的文献资料, 一般采用《中国图书馆图书分类法》进行分类。

3. 主题词途径

主题词是指表征文献内容特征、经过规范化的名词术语。检索者通过检索这些规范的词语来找到所要的文献和情报。(见《现代汉语主题词表》)

4. 关键词途径

直接从文献的篇名或文摘或全文中抽取出来的非规范化检索词来进行检索。

5. 顺序号途径

专利号, 标准号, 研究报告号。

6. 全文检索途径

以一次文献的全部信息作为存储和检索对象的方法。

7. 引文检索途径

利用被引文献来查找原文献的检索方法。如: CNKI中国期刊网被引用字段, 维普中文期刊数据库引文检索途径。

8. 其他途径

如: 超文本检索, 基于概念的自然语言检索, 图像检索等。

1.1.7 文献检索的方法

1. 追溯法

又称回溯法, 是从已有的文献后面所附的参考文献入手, 逐一查找全文, 再从这些原文后面所列的参考文献逐一追查, 不断扩大检索线索, 从而获得一批相关文献信息的查找方法。

2. 工具法 (常用法)

利用文摘、题录或索引等检索工具来查找文献的方法。按所查文献的顺序, 可分为顺查法、倒查法和抽查法三种。

顺查法——由远而近, 从问题发生的年代开始顺着时间顺序从远到近查找, 适用于

无综述性文献可参考时使用。查找的文献较完整,查全率较高,但工作量大,效率不高。

倒查法——由近到远,重点放在近期的文献,多用于新课题或老技术的新发展,从最新情况开始查到一定时间的基本资料为止。

抽查法——根据课题的特点和需要,选查某一时期的文献,可节约时间,但可能会漏检。

3. 综合法

结合追溯法和常用法,先利用检索工具查出一定时期内的一批有用文献,再利用这些文献后面的参考文献,追溯查出前一时期的文献,如此循环交替地使用两种方法直到满足要求为止。也称循环法。

1.1.8 查全率与查准率

查全率和查准率是判定检索效果的主要标准,两者结合起来,描述了系统的检索成功率。

1. 查全率

查全率是指系统在进行某一检索时,检出的相关文献量与系统文献库中相关文献总量的比率,它反映该系统文献库中实有的相关文献量在多大程度上被检索出来。

$$\text{查全率} = [\text{检出相关文献量} / \text{文献库内相关文献总量}] \times 100\%$$

例如,要利用某个检索系统查某课题,假设在该系统文献库中共有相关文献为40篇,而只检索出30篇,那么查全率就等于75%。

2. 查准率

查准率是指系统在进行某一检索时,检出的相关文献量与检出文献总量的比率,它反映每次从该系统文献库中实际检出的全部文献中有多少是相关的。

$$\text{查准率} = [\text{检出相关文献量} / \text{检出文献总量}] \times 100\%$$

如果检出的文献总篇数为50篇,经审查确定其中与项目相关的只有40篇,另外10篇与该课题无关。那么,这次检索的查准率就等于80%。显然,查准率是用来描述系统拒绝不相关文献的能力,有人也称查准率为“相关率”。

3. 影响检索效果的因素

查全率与查准率是评价检索效果的两项重要指标。查全率和查准率与文献的存储与信息检索两个方面是直接相关的,也就是说,与系统的收录范围、索引语言、标引工作和检索工作等有着非常密切的关系。

(1) 影响查全率的因素:检索时检索策略过于简单;选词和进行逻辑组配不当;检索途径和方法太少;检索人员业务不熟练和缺乏耐心;检索系统不具备截词功能和反馈功能,检索时不能全面地描述检索要求等。

(2) 影响查准率的因素:索引词不能准确描述文献主题和检索要求;组配规则不严密;选词及词间关系不正确;标引过于详尽;组配错误。

检索时所用检索词(或检索式)专指度不够,检索面宽于检索要求;检索系统不具备逻辑“非”功能和反馈功能;检索式中允许容纳的词数量有限;截词部位不当,检索式中使用逻辑“或”不当等等。

实际上,影响检索效果的因素是非常复杂的。实验表明,查全率与查准率是呈反比。

关系的。要想做到查全，势必要对检索范围和限制逐步放宽，则结果是会把很多不相关的文献也带进来，影响了查准率。企图使查全率和查准率都同时提高并不是很容易。强调一方面，忽视另一方面，也是不妥当的。应当根据具体课题的要求，合理调节查全率和查准率，保证检索效果。

思考与练习

1. 文献的三要素；信息检索的含义。
2. 什么是顺查法 / 倒查法？
3. 比较题录格式与文摘格式的不同点。
4. 举例说明怎样提高查全率 / 查准率。
5. 举例说明什么是事实检索。