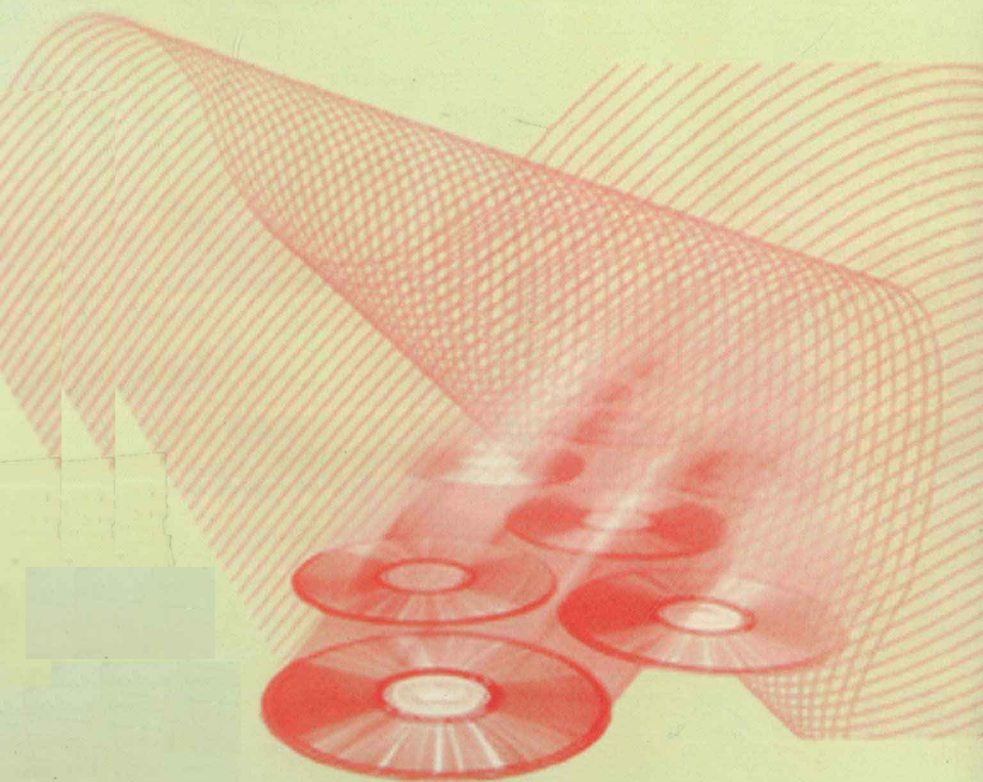


信息检索教程

主编 包和平 宛文红

主审 蔡明德



内蒙古科学技术出版社

信息检索教程

主 编 包和平 宛文红
副主编 张英福 刘江龙
主 审 蔡明德

内蒙古科学技术出版社

图书在版编目(CIP)数据

信息检索教程/包和平,宛文红主编.—赤峰:内蒙古
科学技术出版社,2001.4

ISBN 7-5380-0875-6

I.信… II.①包…②宛… III.情报检索—高等
学校教材 IV.O252.7

中国版本图书馆 CIP 数据核字(2001)第 11445 号

地 址/赤峰市红山区哈达街南一段 4 号

电 话/(0476)8224848 8231924

邮 编/024000

责任编辑/牧仁

封面设计/姜华

印 刷/赤峰地质宏达印刷有限责任公司

开 本/850×1168 1/32

字 数/317 千

印 张/13

印 数/1-3000 册

版 次/2001 年 4 月第 1 版

印 次/2001 年 4 月第 1 次印刷

定 价/24.80 元

前 言

近几十年来,信息检索已发展成为一门具有丰富理论的学科,其中基础理论和方法已发展得相当成熟,并已在实际中广泛应用。特别是高等学校普遍开设信息检索课以来,在广大信息检索课教师的辛勤努力下,信息检索已逐渐成为一门受师生欢迎的基础课,并在高等学校占有不可缺少的一席之地。实践证明,开设信息检索课,对于增强大学生的信息意识,提高信息检索的效率,培养自学能力和独立研究的能力,适应信息化社会更加广泛的信息检索及应用均具有重要的现实意义。

随着信息技术的高速发展,与信息技术密切相关的计算机检索、Internet 国际互联网迅速崛起,对手工信息检索造成极大的冲击,以往的教材已远远不能适应当前教学的需要。面对这一实际情况,我们根据辽宁省教育厅《关于印发辽宁省普通高等学校〈信息检索与利用〉课教学大纲的通知》(辽教办发[2000]87号)精神,结合实际需要,组织信息检索课教师编写了这本《信息检索教程》,旨在为高等学校信息检索课教学尽一份绵薄之力。

本书除阐述了各种信息资源、检索工具书以及大学生必须掌握的手检知识与技能外,对数据库检索、联机检索与 Internet 网络检索作了重点介绍,并在手检与机检相

结合方面进行了有益的探索。本书内容丰富,理论联系实际,重在实践,力求体现学以致用原则和信息检索所具有的方法学特点,同时注意吸收最新的信息检索理论研究成果,追求具有跨世纪的学术前瞻性和应用指导性,以满足新世纪信息检索课教学的需要。

本书共分九章,各章节编写人:第一章为包和平,第二章、第四章为宛文红,第三章为王学艳,第五章为王波,第六章为刘江龙,第七章为刘敏,第八章为张英福,第九章第一、二节为崔亚虹,第三节为高娃,附录由宛文红、金英玉编辑。

本书由包和平、宛文红主编,负责拟总纲及统一全书文字工作。在编写过程中,参考了有关论著,并在书后参考文献中一一列出。对本书所证引资料的原著者,我们表示深切的谢意。因时间匆忙,水平有限,疏漏和欠妥之处在所难免,敬请读者批评指正。

本书的编写和出版,得到了大连民族学院副院长蔡明德研究员、教务处于纯海教授、科研处戴向平副教授的热情支持,在此一并表示衷心地感谢。

目 录

第1章 信息与信息检索概论	(1)
1.1 信息	(1)
1.1.1 信息的定义	(1)
1.1.2 信息的特征	(3)
1.1.3 信息的分类	(4)
1.2 信息资源	(4)
1.2.1 信息资源的类型	(5)
1.2.2 文献信息资源	(5)
1.2.3 电子信息资源	(9)
1.3 信息检索.....	(12)
1.3.1 什么是信息检索.....	(13)
1.3.2 信息检索的作用.....	(14)
1.3.3 信息检索技术的发展过程.....	(15)
第2章 信息检索的基本原理与方法	(18)
2.1 信息检索语言	(18)
2.1.1 信息检索语言的定义及功能	(18)
2.1.2 信息检索语言的分类	(19)
2.1.3 等级体系分类语言.....	(20)
2.1.4 主题法语言.....	(31)
2.2 信息检索系统.....	(41)
2.2.1 信息检索系统的有机组成.....	(42)
2.2.2 信息检索系统的检索效率.....	(43)
2.3 检索工具.....	(46)

2.3.1	检索工具的基本职能	(46)
2.3.2	检索工具的类型	(47)
2.4	检索途径	(51)
2.5	检索方法	(53)
2.6	检索策略	(54)
2.6.1	什么是检索策略	(54)
2.6.2	构造检索策略的步骤和方法	(55)
第3章	中文工具书及其使用方法	(59)
3.1	中文工具书概述	(59)
3.1.1	中文工具书的历史、现状及发展趋势	(59)
3.1.2	中文工具书的特点及分类	(61)
3.1.3	中文工具书排检法	(63)
3.2	中文检索工具书及其使用方法	(66)
3.2.1	中文检索工具书体系	(66)
3.2.2	常用中文检索工具	(69)
3.2.3	中文检索工具内容编排与著录格式	(80)
3.2.4	检索途径与检索示例	(83)
3.3	参考工具书及其使用方法	(86)
3.3.1	参考工具书结构与编排	(86)
3.3.2	参考工具书类型	(88)
第4章	国外著名检索工具及其使用方法	(99)
4.1	国外检索工具简介	(99)
4.2	美国《工程索引》	(99)
4.2.1	概况	(99)
4.2.2	Ei 的出版形式	(100)
4.2.3	Ei 正文的编排形式及著录格式	(102)
4.2.4	Ei 的辅助索引和附表	(103)
4.2.5	Ei 的检索方法	(107)

4.3	英国《科学文摘》	(109)
4.3.1	概况	(109)
4.3.2	SA 正文的编排形式及著录格式	(110)
4.3.3	SA 的辅助索引和附表	(112)
4.3.4	SA 的检索方法	(116)
4.4	美国《化学文摘》	(116)
4.4.1	概况	(116)
4.4.2	CA 正文的编排形式及著录格式	(117)
4.4.3	CA 的辅助索引	(120)
4.4.4	CA 的检索方法	(127)
4.5	《科学引文索引》	(129)
4.5.1	概况	(129)
4.5.2	SCI 的编排体例和著录格式	(130)
4.5.3	SCI 的检索方法	(134)
4.6	俄罗斯《文摘杂志》(PЖ)	(135)
4.6.1	概况	(135)
4.6.2	PЖ 正文的编排体例和著录格式	(136)
4.6.3	PЖ 的辅助索引和附表	(138)
4.6.4	PЖ 的检索方法	(140)
4.7	日本《科学技术文献速报》	(141)
4.7.1	概况	(141)
4.7.2	《速报》的编排体例和著录格式	(142)
4.7.3	《速报》的辅助索引和附表	(143)
4.7.4	《速报》的检索方法	(145)
第5章	特种文献及其检索方法	(146)
5.1	科技报告及其检索方法	(146)
5.1.1	概述	(146)
5.1.2	PB 报告概况	(146)

5.1.3	AD 报告概况	(147)
5.1.4	PB、AD 报告的检索工具—— 美国《政府报告通报和索引》	(148)
5.1.5	NASA 报告及其检索工具	(150)
5.1.6	AEC/ERDA/DOE 报告及其检索	(151)
5.2	会议文献及其检索	(153)
5.2.1	概述	(153)
5.2.2	国外会议文献的检索工具	(154)
5.2.3	国内会议文献的检索工具	(159)
5.3	专利文献及其检索	(160)
5.3.1	专利及专利制度	(160)
5.3.2	专利文献	(161)
5.3.3	国际专利分类法	(163)
5.3.4	德温特专利及其检索	(165)
5.3.5	《WPI 优先案对照索引》	(169)
5.3.6	美国专利检索工具——《专利公报》	(170)
5.3.7	中国专利文献检索	(170)
5.4	标准文献及其检索	(173)
5.4.1	概述	(173)
5.4.2	国外标准文献及其检索	(174)
5.4.3	国内标准文献及其检索	(177)
5.5	学位论文、产品样本及其检索	(179)
5.5.1	学位论文及其检索	(179)
5.5.2	产品样本及其检索	(182)
第 6 章	信息数据库检索	(185)
6.1	数据库发展概况	(185)
6.1.1	国外数据库的发展	(185)
6.1.2	国外数据库特征	(187)

6.1.3	国内数据库概况	(191)
6.1.4	国内数据库与国外数据库的差距	(194)
6.1.5	数据库产业的发展趋势	(195)
6.2	文摘数据库的检索方法	(197)
6.2.1	信息数据库的原理	(197)
6.2.2	《中文科技期刊数据库光盘》检索	(199)
6.2.3	《中国专利数据库》检索	(204)
6.2.4	《美国商务信息数据库》检索	(216)
6.3	全文数据库检索方法	(220)
6.3.1	《中国学术期刊》全文数据库收藏的条件	(220)
6.3.2	CAJCD 的检索特点	(221)
6.3.3	CAJCD 的检索方法	(221)
第7章	国际联机信息检索	(230)
7.1	数据库	(230)
7.1.1	数据库属性	(230)
7.1.2	数据库字段和记录	(232)
7.1.3	数据库索引和数据库标引词典	(233)
7.1.4	数据库类型	(233)
7.2	联机信息检索	(233)
7.2.1	联机信息检索特点	(233)
7.2.2	联机信息检索类型	(235)
7.2.3	联机信息检索操作步骤	(236)
7.2.4	著名联机信息检索系统简介	(236)
7.3	联机信息检索软件与检索方式	(240)
7.3.1	联机信息检索软件	(240)
7.3.2	联机信息检索系统的功能和特性	(242)
7.3.3	联机信息检索方式	(245)
7.3.4	Sci search 检索方法	(262)

7.3.5	Ei Compendex Plus 检索方法	(264)
7.4	联机信息检索系统与 Internet	(265)
7.4.1	DIALOG 系统的 Internet Web 界面的检索方法	(265)
7.4.2	OCLC 第一检索服务(First Search Web)	(266)
第8章	Internet 网络信息检索	(267)
8.1	Internet 简介	(267)
8.1.1	什么是 Internet	(267)
8.1.2	Internet 的起源和发展	(268)
8.1.3	中国的 Internet	(269)
8.1.4	Internet 常用术语	(275)
8.1.5	Internet 的主要服务	(279)
8.2	Internet 信息检索工具	(280)
8.2.1	网络信息检索工具的基本原理与方法	(281)
8.2.2	网络信息检索工具类型	(286)
8.2.3	英文综合型网络信息检索工具	(293)
8.2.4	中文综合型网络信息检索工具	(307)
8.3	网络电子信息源及其检索方法	(316)
8.3.1	电子图书检索	(317)
8.3.2	电子期刊检索	(319)
8.3.3	电子报纸检索	(322)
8.3.4	数字图书馆	(324)
第9章	科技论文的写作	(330)
9.1	论文的分类	(330)
9.1.1	从内容上划分	(330)
9.1.2	从写作目的上划分	(331)
9.1.3	从研究方法上划分	(335)
9.1.4	从论证性质的角度划分	(337)

9.1.5 从论述范围的角度划分	(338)
9.2 科技论文的写作要求	(338)
9.2.1 引言的撰写	(339)
9.2.2 正文的撰写	(342)
9.2.3 结论的撰写	(345)
9.3 科技论文的写作格式	(347)
9.3.1 前置部分	(347)
9.3.2 主体部分	(354)
9.3.3 后置部分	(354)
附录 1 中国六大主干网及主要教育、科研机构网址	(357)
附录 2 中国图书馆网址	(362)
附录 3 世界主要图书馆网址	(369)
附录 4 中外文搜索引擎目录	(371)
附录 5 中国期刊网址	(376)
附录 6 世界主要期刊网址	(378)
附录 7 中国报纸网址	(383)
附录 8 世界主要报纸网址	(387)
附录 9 国外 100 所著名大学网址	(390)
附录 10 国际联机检索提问单	(398)
附录 11 DIALOG 系统主要数据库收费标准	(400)
参考文献	(404)

第 1 章 信息与信息检索概论

1.1 信息

世纪之交,人类社会已经感受到信息时代的强有力冲击,信息资源以势不可挡的影响力成为维系人类社会生存和发展的三大资源之一(与物质和能量并列)。可以说,信息已经与每一个人息息相关。人们只有充分认识信息,学会捕捉和利用信息,才能在信息时代站稳脚跟,掌握住生存之道。

1.1.1 信息的定义

信息一词的拉丁词源是 Information,意思是通知、报道或消息。在中国历史资料中,信息一词最早出自唐诗,是音信、消息的意思。一直沿用到 20 世纪中叶,其科学含义才被逐渐揭示出来。到目前为止,关于信息的定义已不下一百种,它们都从不同的侧面揭示了信息的某些属性和特征,但据科学界分析,至今尚无一种定义,因其内涵的全面性、深刻性和科学性而被社会各界一致认同。但许多定义仍不乏一定的探索意义,这里引述几种较有代表性和影响力的说法供读者参考,希望读者能从中获取进一步探索的灵感。

(1) 信息是一切事物的存在方式、运动状态及其表征。

这种定义包含两层含义:从本体论意义上说,信息泛指一切事物运动的状态和运动的方式,包括事物内部结构的状态和方式,以及外部联系的状态和方式;从认识论意义上说,信息是关于事物运

动的状态和方式的反映。

(2) 信息是用来消除不定性的东西。

正因为信息是事物运动的状态和方式(本体论意义),是关于事物运动状态和方式的表征(认识论意义),它才可以用来消除人们认识上相应的不定性。这是以上两种定义之间的内在联系。

(3) 信息是一切事物发出的消息、指令、数据、符号等所包含的内容。

事实上,任何一种消息、指令等都不外是关于某种事物的运动状态和运动方式的某种形式的反映。因此,这个定义与认识论意义上的信息含义也是相通的。

(4) 信息是事物联系的普遍形式。

人们通过获得、识别自然界和社会的不同信息来区别不同事物,得以认识和改造世界。在一切通讯和控制系统中,信息是一种普遍联系的形式。

(5) 信息是通讯系统传输和处理的内容。

(6) 信息是指应用文字、数据和信号等形式,通过一定的传递和处理,来表现各种相互联系的客观事物在运动变化中所具有的特征性内容的总称。

从这几则定义我们可以得出以下结论:信息是普遍存在于自然界和社会界中的一切事物存在和变化的特征性反映;是事物相互作用和相互联系的表征;是人类感知和未感知到的事物的普遍属性。

对人类和人类社会而言,不论信息的真实含义是否被充分认识,它对人类都同样有着特殊的重要意义,和物质、能量一样,它是一种重要的资源。物质资源提供各种各样的材料,能量资源提供形形色色的动力,而信息提供的是知识和智慧。人类只有不断地从外界取得信息,并加工、积累成知识,才能有效地改造世界。因此没有信息,就不能认识世界,更谈不上改造世界。从这个意义

上,可以说信息是认识的向导、实践的指南、决策的依据、控制的基础、系统的灵魂。

时值今日,随着科学技术日新月异的发展,信息已经与计算机技术、通讯技术、网络技术及多媒体、超文本等概念紧密地联系起来,我们应该不断从认识论和实践论的角度拓展对信息的理解,将因时代的变革而赋予它的新含义揭示出来,以便更好地利用它来改造世界。

1.1.2 信息的特征

信息作为一种新型的重要资源,在其被摄取和使用过程中表现出了区别于物质和能量的特征。

(1) 相对独立性

信息来源于事物(物质),而不是事物本身,它可以脱离母体而相对地独立。正如前面所述,信息是“事物的运动状态、存在方式及其表征”,一旦被观察者所感知、获取,便可以脱离母体而存在。这样,人们就可以通过信息对事物进行间接的分析和研究。否则,人们必须亲自接触才能认识某种事物。

(2) 转移性(传递性)

由于信息具有脱离母体而相对独立的能力,因而它可以通过一定的方法在时间上或在空间中进行转移,或者说传递。也正因为这种可转移的特征,人们才得以对信息进行存储、传播和处理。

(3) 变换性

信息在转移过程中,其载体形态是可变的,它可以由一种载体转换成另一种载体。信息既然不是事物本身,它就可以负载在其他一切可能的载体形态上,例如数据、文字、图像等等。

(4) 共享性

信息可以被不同用户所共享。基于信息可脱离母体而独立存在于不同载体的特性,信息可以被无限地复制和传递,而且任何

一个信息拥有者不会因把信息传递出去而丧失对该信息的拥有，人们可以同时拥有同样的信息。

信息的这几种主要特性，决定了它可以被人类及其它生物等所摄取、变换、传递、存储、识别、加工、检索及利用等。

1.1.3 信息的分类

(1) 按信息的生成领域划分

可分为社会信息、自然信息两大类。

(2) 按信息的应用方向划分

分为科技信息、政治信息、经济信息、文化信息、工农业生产信息、生活信息等。

(3) 按信息加工程度划分

包括一次信息、二次信息和三次信息。一次信息，通常指未经任何加工的原始信息。二次信息，是对一次信息进行收集、整理和加工，并按一定的方法编排而成的信息，如书目、索引、文摘等。二次信息是查找和利用一次信息的重要工具，是信息检索理论研究的核心内容。三次信息是在二次信息的基础上，对一次信息进行高度概括、论述、分析、综合后形成的参考价值很高的信息。如综述、专题报告等。

(4) 按信息的载体形态划分

分为文字信息、数据信息、图像信息、多媒体信息等等。

(5) 按信息的传播范围划分

分为公开信息、半公开信息、非公开信息等等。

1.2 信息资源

如前所述，普遍存在于自然界和人类社会的信息已经成为与物质和能量并列的，维系人类社会生存的三大资源之一。信息既

是一种重要的自然资源,又是不可忽视的社会资源。它是包含信息本身、信息生产者、信息加工和整理人员、信息技术设施及信息机构等多种相关因素在内的信息资源系统。这是对信息资源的广义理解。狭义的理解是,信息源于世间万物,而不是事物本身,具有脱离事物相对独立的特性,因此信息必须借助语言、文字、图像、数据等一切可能的形式存储在各种介质的载体上被传播、处理和利用。而存有信息的各种载体的集合就是信息资源的狭义内涵,亦称信息源。下面着重介绍狭义内涵的信息资源。

1.2.1 信息资源的类型

信息资源的类型依不同标准划分如下:

(1) 按信息内容的记录形式,可分为文字型信息源、数据型信息源、声像型信息源和多媒体信息源。

多媒体信息源是集文字、声音、图像于一体的现代化新型信息源,多以光盘和网络资源等形式出现,有较大发展潜力。

(2) 按信息的载体介质划分,主要有印刷型信息源、电子信息源。

(3) 按记录信息的出版形式划分,有图书、报纸、期刊、科技报告、会议录、学位论文、专利文献、标准文献、产品样品等。

(4) 按信息内容的应用领域划分,包括科技、经济、管理、生活等信息源。

1.2.2 文献信息资源

所谓文献,是记录有知识的一切载体。凡是人类用文字、符号、图像、声音等记录下来的东西,均称之为文献。因此,文献一直是最常用,也是最丰富的一种信息资源。到目前为止,文献主要包括以下几种类型:

(1) 不同加工程度的文献