

21

世纪高职高专规划教材

计算机信息检索与利用

主编 吕润桃 覃国萍 副主编 李 璞 刘 伟 主审 禹文龙

21SHJGGAOZHIGGAOZHUAJIAOCA



中国水利水电出版社

www.waterpub.com.cn

21 世纪高职高专规划教材

计算机信息检索与利用

主 编 吕润桃 覃国萍

副主编 李 瑛 刘 伟

主 审 禹文龙



中国水利水电出版社
www.waterpub.com.cn

内 容 提 要

本书的目的是为在网上查询信息的用户提供搜索工具的基本知识。本书共有 6 章, 介绍网络信息检索的理论基础和一般方法, Internet 信息资源检索与利用、其他信息资源检索与利用、中文数据库资源检索与利用、英文数据库资源检索与利用及科技论文写作。

本书可作为高职高专学校各专业信息检索课程的教材, 也可作为各类图书馆和信息机构的岗位培训教材和参考书。

图书在版编目 (CIP) 数据

计算机信息检索与利用 / 吕润桃, 覃国萍主编. —北京:
中国水利水电出版社, 2009
21 世纪高职高专规划教材
ISBN 978-7-5084-6541-8

I. 计… II. ①吕…②覃… III. 计算机应用—情报检索—
高等学校: 技术学校—教材 IV. G252.7

中国版本图书馆 CIP 数据核字 (2009) 第 132610 号

策划编辑: 杨庆川 责任编辑: 宋俊娥 加工编辑: 谢 琛 封面设计: 李 佳

书 名	21 世纪高职高专规划教材 计算机信息检索与利用
作 者	主 编 吕润桃 覃国萍 副主编 李 瑛 刘 伟 主 审 禹文龙
出版发行	中国水利水电出版社 (北京市海淀区玉渊潭南路 1 号 D 座 100038) 网址: www.waterpub.com.cn E-mail: mchannel@263.net (万水) sales@waterpub.com.cn 电话: (010) 68367658 (营销中心)、82562819 (万水)
经 售	全国各地新华书店和相关出版物销售网点
排 版	北京万水电子信息有限公司
印 刷	北京市天竺颖华印刷厂
规 格	184mm×260mm 16 开本 10 印张 221 千字
版 次	2009 年 8 月第 1 版 2009 年 8 月第 1 次印刷
印 数	0001—5000 册
定 价	18.00 元

凡购买我社图书, 如有缺页、倒页、脱页的, 本社营销中心负责调换
版权所有·侵权必究

前 言

当今时代是一个信息时代，信息对于经济和社会的发展、科技文化的进步都起着重要的作用。而计算机的普及更是锦上添花，20 世纪 90 年代以来，因特网在我国迅猛发展，网络信息的飞速增长极大地改变了人们获取信息的方式，因特网已经成为我们学习生活不可或缺的重要组成部分。人们上网的主要目的是获取信息。网络已经改变了人们使用图书馆的方式，如果要查找资料，大部分人会选择上网。

在网络化和数字化时代，传统的信息检索理论和方法已经出现了很大的变化，数字化技术、全文检索技术、因特网技术、多媒体技术和软硬件的频繁更新换代，使得即使有一定水平的专业人员也不能只停滞在原有的知识层面上，而是需要不断更换新的工具，采用新的方法，获取新的知识和技能。因此，信息的科学检索和利用显得尤其重要。针对这样的现状，本书在参考其他同类最新书籍和网页信息的基础上，主要介绍信息检索基础知识和常用的网络信息资源及其检索等内容。通过对几种常用的中外数据库的基本检索方法及网络信息检索技巧等的重点介绍，旨在提高网络环境下信息资源的利用能力。在此基础上，增加了网络信息利用、梳理、整合等相关方面的内容，如资料的加工与整理、科技论文的写作方法等。本书图文并茂、例证丰富、内容明晰，可作为高职高专院校的公共课教材，也可作为广大教师、科研人员、岗位培训等信息检索的参考用书。

本书共 6 章，介绍计算机信息检索的基础知识、Internet 信息资源检索与利用、其他信息资源检索与利用、中文数据库资源检索与利用、英文数据库资源检索与利用、科技论文写作及检索。

本书由包头轻工职业技术学院吕润桃、覃国萍任主编，李瑛、刘伟任副主编，禹文龙主审。参与本书编写的人员还有宋耀东、韩耀坤、郭洪兵、孙元、张尼奇、刘涛、王慧敏、孙妍、赵志茹、张英芬、刘媛、刘婧婧、武婧伟、木其乐。

需要申明一点，网络上的信息“新陈代谢”速度极快，部分网站会发生变动，有的关闭，有的更新，敬请读者注意，如果发现网站内容变化，可用搜索引擎寻找更新后的地址。书中不免有疏漏之处，敬请读者批评指正。

编 者

2009 年 6 月

目 录

前言

第 1 章 计算机信息检索基础知识	1
1.1 信息检索概述	1
1.1.1 信息的定义	1
1.1.2 信息的特征	1
1.1.3 文献的定义	2
1.1.4 文献的分类	2
1.1.5 信息检索的含义	3
1.1.6 信息检索方式	4
1.1.7 信息检索工具	4
1.1.8 信息检索工具的类型	4
1.1.9 信息检索的类型	5
1.1.10 信息检索语言	6
1.1.11 信息检索的常用途径	8
1.1.12 信息检索的方法	9
1.2 计算机信息检索基础	9
1.2.1 计算机检索概述	9
1.2.2 计算机检索的含义	10
1.2.3 计算机检索的特点	11
1.2.4 计算机信息检索的原理	11
1.2.5 计算机信息检索系统的构成	11
1.2.6 计算机信息检索的基本技术	13
1.2.7 计算机检索的技巧和方法	14
1.2.8 查全率与查准率	15
1.2.9 信息的鉴别与筛选	16
1.2.10 信息的加工与整理	16
习题一	16
第 2 章 Internet 信息资源及其检索	17
2.1 Internet 信息资源及其组织方式	17
2.1.1 Internet 信息资源的定义	17
2.1.2 Internet 信息资源的特点	17
2.1.3 Internet 信息资源及检索方法	18

2.2	Internet 信息资源检索工具	19
2.2.1	搜索引擎	19
2.2.2	搜索引擎的搜索方式	20
2.2.3	阅读软件的安装和使用	28
2.3	常用中文搜索引擎	31
2.3.1	谷歌搜索引擎	31
2.3.2	百度搜索引擎	36
2.3.3	雅虎搜索引擎	37
2.3.4	搜狐搜索引擎	39
2.3.5	天网搜索引擎	40
2.3.6	CNKI 搜索引擎	40
2.3.7	元搜索引擎	41
2.3.8	其他的搜索引擎	42
	习题二	42
第 3 章	中文数据库资源检索与利用	43
3.1	CNKI 数据库系统	44
3.1.1	初级检索	45
3.1.2	高级检索	52
3.1.3	专业检索	55
3.2	万方数据资源系统	58
3.2.1	万方数据资源子系统	58
3.2.2	快速检索	61
3.2.3	跨库检索	64
3.3	维普资讯网	65
3.3.1	文章搜索	66
3.3.2	分类导航	70
3.3.3	期刊导航	73
3.4	超星数字图书	74
3.5	书生数字图书	78
3.6	龙源期刊网	79
3.6.1	期刊	79
3.6.2	电子期刊	80
3.6.3	龙源期刊网	80
3.7	其他中文数据库	82
3.7.1	国家科技图书文献中心数据库	82
3.7.2	CALIS 中文资源数据库	82
3.7.3	CASHL 资源与服务	83

3.7.4	免费中文科技文献数据库	84
3.7.5	国内上网图书馆	85
3.7.6	特种文献数据库	86
3.7.7	其他中文数字图书馆	89
3.7.8	中文科学引文索引	89
习题三		90
第4章	其他科技信息资源的检索与利用	91
4.1	专利文献检索	91
4.1.1	基本概念	91
4.1.2	中华人民共和国国家知识产权局	92
4.1.3	中国专利信息网	95
4.2	标准文献检索	96
4.2.1	基本概念	96
4.2.2	中国标准服务网	97
4.2.3	中国标准化协会信息网	98
4.2.4	国际标准化组织 ISO 网	99
4.2.5	标准全文的索取	99
4.3	科技报告	99
4.3.1	概述	99
4.3.2	国外有名的科技报告	100
4.3.3	科技报告的原文索取	100
4.4	会议文献	101
4.4.1	概述	101
4.4.2	国内外会议文献检索	101
4.5	学位论文	102
4.5.1	概述	102
4.5.2	学位论文检索	102
4.6	工具书	102
4.6.1	概述	102
4.6.2	工具书种类	103
4.7	技术档案	104
4.8	产品资料	104
4.9	国内其他科技信息网站	105
习题四		107
第5章	英文数据库资源检索与利用	108
5.1	参考数据库概述	108
5.2	常用参考数据库	108

5.2.1	《科学引文索引 (SCI)》	108
5.2.2	《社会科学引文索引 (SSCI)》	109
5.2.3	《艺术与人文科学引文索引 (A&HCI)》	110
5.2.4	《化学文摘 (CA)》	110
5.2.5	《科学技术会议录索引 (ISTP)》	110
5.2.6	《工程索引 (EI)》	110
5.3	全文数据库概述	119
5.4	常用全文数据库	119
5.4.1	EBSCO 全文数据库	119
5.4.2	Springer 电子期刊数据库	121
5.4.3	Elsevier (ScienceDirect) 数据库	122
习题五		123
第 6 章	科技论文写作及其检索	125
6.1	科技论文的定义与特点	125
6.1.1	科技论文的定义	125
6.1.2	科技论文的特点	125
6.2	科技文章分类与各类论文的写作特点	126
6.2.1	科技文章的种类	126
6.2.2	常见科技论文的写作要求和特点	126
6.3	科技论文写作国家标准	129
6.4	科技论文写作	131
6.4.1	论文选题	131
6.4.2	开题报告	131
6.4.3	科技论文的内容	132
6.4.4	论文的修改	143
6.5	科技查新	143
6.5.1	科技查新定义	143
6.5.2	科技查新对象	144
6.5.3	科技查新作用	144
6.5.4	科技查新的结果	144
6.5.5	科技查新步骤	145
6.5.6	科技查新报告	146
6.5.7	附件	148
习题六		148
参考文献		149

第 1 章 计算机信息检索基础知识

1.1 信息检索概述

1.1.1 信息的定义

信息是物质的一种基本属性。它是自然和人类社会中一切事物自身的运动状态与方式，是事物的内在联系和含义的表征。

1.1.2 信息的特征

1. 客观性

信息不是虚无缥缈的事物，它的存在可以被人们感知、获取、传递和利用。信息是现实世界中各种事物运动与状态的反映，其存在是不以人的意志为转移的。客观、真实是信息的最重要的本质特征。

2. 时效性

由于事物是在不断变化着的，那么表征事物存在方式和运动状态的信息也必然会随之改变。在现代社会中，信息的使用周期越来越短，信息的价值实现取决于对其及时地把握和运用。如果不能及时地利用最新信息，信息的价值就会贬值甚至毫无价值，这就是信息的时效性，即时间与效能的统一性。它既表明了信息的时间价值，也表明了信息的经济价值。

3. 载体性

信息必须依附于一定的载体（如声波、电磁波、纸张、化学材料、磁性材料等）才能流通和传递，否则，信息的价值就不能体现。信息可以存储在不同的载体上，但其内容并不因记录手段或物质载体的改变而发生变化。例如关于筑路机械的信息，不论是刊登在报刊上、发布在电视节目中还是存储在光盘数据库中，其信息内容和价值是同样的。

4. 传递性

信息依附于一定的物质载体后，其传递和流通便成为可能。信息的传递性是指信息从信源出发，经过信息载体的传递被信宿接收并进行处理和利用的特性。不同载体的信息可以通过计算机、人际交流、文献交流或大众传媒等手段传递给信息用户，这种跨越时空的传递特性是实现信息资源共享的基础，是将信息最大化利用的保证。

5. 可塑性

信息在流通和使用过程中，人们借助于先进的技术，可以对其进行综合、分析及加工处理。也就是把信息从一种形式变换成另一种形式，如可以将一本图书加工为题录或文摘等形式，从而方便用户的选择和利用。不过，在信息的加工过程中，信息量会减少或增加。

用户可根据检索需要选择不同的信息形式。

6. 共享性

共享性是指同一信息同时或不同时被多个用户使用,而信息的提供者并不因此而失去信息内容和信息量。信息的共享性可以提高信息的利用率,人们可以利用他人的研究成果进一步创造,避免重复研究,节约资源。

1.1.3 文献的定义

文献是记录有知识的一切载体的统称,即用文字、图像、符号、声音、音频、视频、编码等手段记录的某种知识信息的各种载体,具有存储和传递知识信息的功能。

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。

1.1.4 文献的分类

文献按其出版形式分类为:

(1) 图书(Book)。图书是历史最悠久、最传统的信息类型。其内容是总结性的,具有全面、系统、成熟的特点,但出版周期长,传递信息的速度较慢,老化速度快。近年来增长迅速的电子书可在一定程度上弥补印刷型图书的不足。

(2) 期刊(Periodical)。期刊比图书的出版周期短,刊载文献速度快、数量多、内容新,能及时反映某一研究领域的学术发展水平,所以很受用户的重视。随着电子信息资源的大量涌现,信息用户将主要通过期刊全文数据库实现检索目的。

(3) 会议论文(Conference Paper)。会议论文又称会议文献,是将学术会议或专业会议上交流的论文编辑出版的信息。会议论文内容新颖、专业性和针对性强,传递信息迅速,能及时反映某个专业领域的研究水平和动向,是了解学科发展的一种重要的信息源。但未经出版者公开发行的会议文献一般难以获得。

(4) 专利文献(Patent Document)。专利文献又称专利说明书,是一种集经济、技术与法律为一体的信息形式。它是专利申请人向专利机构提交的说明该项发明的目的、优点、技术原理和专利权限的书面文件。也就是说,专利文献是发明的书面形式。一项新的专利代表着一个科研领域的最高水平。毋庸置疑,专利说明书包含丰富的技术情报,有极高的含金量。据统计表明,全世界新技术的90%~95%是通过专利文献公布于世的。专利说明书主要由专利文献馆和省级以上科技信息所收藏。

(5) 科技报告(Sci-Tech Report)。科技报告又称为技术报告、研究报告。有些科技报告是由政府部门出版的,所以也称之为政府报告。技术报告是记录研究工作和开发调查工作的成果或进展情况的一种信息类型。科技报告的内容新颖、详尽、专深,出版及时,报道科技成果的速度要快于期刊及其他文献。大部分科技报告是保密或控制发行的,

只有少数可以公开或半公开发表,不易获取原文,因此它又被称为“灰色文献(Grey Document)”。

最著名的科技报告是美国的四大报告:①PB报告(土木建筑、城市规划、环境污染、生物医学等);②AD报告(军事、航空航天、物理、材料工程技术等);③NASA报告(航空和空间科学等);④DOE报告(能源保护、矿物燃料、核能、太阳能等)。在我国,国家图书馆、上海图书馆、中国科技信息研究所和国防科技信息研究所等收藏有较全面的科技报告。

(6) 标准(Standard)。标准主要指技术标准,是对各种产品、元件、工程建设质量、规格、检验方法等作业的技术规定,具有约束性、时效性和针对性,是各类生产建设工作的共同依据。标准对于产品更新换代、工艺水平的改进、产品质量的提高以及市场竞争力的加强,可以起到非常重要的作用。中国标准化综合研究所标准馆、省级技术监督部门的文献馆、科技信息所以及图书馆收藏有标准文献。

(7) 学位论文(Dissertation)。学位论文是高等学校或研究机构的学生为获得某种学位而撰写的科学论文。一般指学士论文、硕士论文和博士论文。不过学位论文的质量参差不齐,其中,博士论文的参考价值较大。学位论文一般不出版发行,通常只在学位授予单位的图书馆和按国家规定接受呈缴本的国家版本图书馆保存有副本。

(8) 技术档案(Technical Records)。技术档案是在技术活动中形成的,是某个工程对象的技术文件、图样、图表、照片或原始记录等。其内容真实、详尽,准确可靠,是科研和生产建设工作的重要依据。它可以重复实践或改进后重复实践,而且具有明显的保密性,具有很大的参考价值。国家级和省级的档案馆是保存档案的重要场所。

除以上介绍的信息类型外,信息的出版形式还包括产品资料、报纸、计算机软件等。

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。

文献按其文献载体形式分类为:

(1) 印刷型文献。

(2) 机读型文献:需要借助一定的机器设备方可阅读的文献。

(3) 缩微型文献。

(4) 视听型文献。

(5) 数字化文献。

印刷型和数字化型是主要的文献载体形式。

1.1.5 信息检索的含义

从广义上来说,信息检索是指将信息按一定的方式组织和存储起来,并根据信息用户

的需要找出有关的信息过程，故全称为信息的存储与检索（Information Storage and Retrieval）。

从狭义上来说，信息检索仅指上述过程的后半部分，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寻（Information Search）。

可见，信息检索的全过程应包括两个主要方面：

（1）信息标引和存储过程。标引是用检索语言和分类号、主题词表示信息，通过对大量无序的信息资源进行标引处理，使之有序化，并按科学的方法存储组成检索工具或检索文档——组织检索系统的过程。

（2）信息的需求分析和检索过程。分析用户的信息需求，利用组织好的检索系统，按照系统提供的检索方法和途径检索有关信息——检索系统的应用过程。

因此，信息检索的实质是将描述用户所需信息的提问特征，与信息存储的检索标识进行比较，从中找出与提问特征一致或基本一致的信息。所谓提问特征就是对信息的需求分析后，从中选出能代表信息需求的主题词、分类号或其他符号。

1.1.6 信息检索方式

1. 手工检索

人们利用图书馆或资料中心，直接查找印刷型文献的检索方式，主要是靠“手翻、眼看、大脑判断”。

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。

1.1.7 信息检索工具

1. 定义

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。

2. 功能

信息检索工具具有报道文献信息、存储文献信息、检索文献信息的功能。

1.1.8 信息检索工具的类型

1. 按信息存储和检索所采用的设备和手段划分

手工信息检索系统分为书本式检索系统和卡片式检索系统。

计算机信息检索系统分为光盘检索系统、联机检索系统和网络检索系统。

2. 按揭示信息内容的程度划分

（1）目录（Catalog）是以完整的出版单元（如一种图书、一种期刊）为单位，按照一定次序编排的对文献信息进行描述和报道的工具，也称为书目。目录对文献的描述比较简单，每条记录的字段主要包括文献题名、责任者、出版事项、分类号、主题词等。一种出版物经过如此描述后形成一条记录，将所有的记录组织起来就形成了目录。

按照收录文献的类型可分为：图书目录、期刊目录、标准目录、报纸目录、地图目录、档案目录等。

按照收录文献的内容范围可分为：综合目录、专题和专科目录、个人著述目录（揭示与报道特定人物的全部著作以及有关该人物的文献）、地方文献目录（揭示与报道某地区文献的目录）等。

按照功用划分，有登记书目、书业书目、收藏目录、专题目录、推荐书目、书目之书目。

（2）题录（Bibliography）是描述文献各种特征的一条记录，通常由文献的分类号、文献篇名、著者及其单位、文献来源、主题词等项组成。题录型检索工具（简称题录）是以单篇文献（如一篇论文、书中的一章）为报道单位。这是它与目录的主要区别，题录揭示文献比目录更为详细。

（3）文摘（Abstract）是指通过描述文献的各种特征，并介绍文献主要内容来报道文献的工具。它与题录型检索工具的报道单位一样，也是单篇文献，只是著录项目（字段）比题录多了一个摘要项。文摘型检索工具（简称文摘）是书目检索系统的核心，是书目检索的主要工具。

（4）索引（Index）是将某种信息集合（如目录、题录、文摘）中的一组相关信息，按照某种顺序组织并指引给用户的一种指南。它具有便于检索，揭示文献信息深入、详细的优点。通常，索引从属于特定的信息集合，为其提供多种不同的检索途径（入口），在检索系统中占有重要地位。索引条目一般由检索标识和存储地址组成。标识是索引条目所指示的文献信息某方面的特征，如主题词、著者姓名、分类号等；存储地址是指明标识所表达的特定信息在信息集合中的地址，它通常是文献存取号（文摘号）。不同的标识系统构成不同的索引，常见的索引有以下几种：

- 分类索引：是以分类号或类目标名称作为索引标识，按照分类号排列形成的索引。
- 主题索引：是以主题词（叙词或关键词）作为索引标识，按其字顺排列形成的索引。
- 著者索引：是以文献上署名的著者、译者、编者等责任者的姓名或机关团体名称作为索引标识，按其字顺排列形成的索引。
- 专用索引：是以某些领域专用的名词术语或符号作为索引标识编排形成的索引。
- 引文索引：是以引文著者和引文的其余题录部分作为标识编制成的索引。

1.1.9 信息检索的类型

1. 按检索内容分，有数据信息检索、事实信息检索和文献信息检索

（1）数据信息检索（Data Information Retrieval）是将经过选择、整理、鉴定的数值数据存入数据库中，根据需要查出可回答某一问题的数据的检索。这些数据包括物理性能常数、统计数据国民生产总值、外汇收支等。这类检索不仅查出数据，还可以提供一定的推导、运算的能力。

（2）事实信息检索（Fact Information Retrieval）是将存储于数据库中的关于某一事件

发生的时间、地点、经过等情况查找出来的检索。它既包含数值数据库的检索、运算、推导,也包括事实、概念等的检索、比较、逻辑判断。

(3) 文献信息检索(Document Information Retrieval)是将存储于数据库中的关于某一主题文献的线索查找出来的检索。它通常通过目录、索引、文摘等二次文献,以原始文献的出处为检索目的,可以向用户提供原文献的信息。一些观点也认为可称为“数目检索”。

2. 按组织方式分,有全文检索、超文本检索和超媒体检索

(1) 全文检索(Full Text Retrieval)是将存储在数据库中的整本书、整篇文章中的任意内容信息查找出来的检索。可以根据需要获得全文中的有关章、节、段、句、词等的信息,也可进行各种统计和分析。

(2) 超文本检索(Hyper Text Retrieval)是对每个节点中所存的信息以及信息链构成的网络中信息的检索。强调中心节点之间的语义联结结构,靠系统提供的工具进行图示穿行和节点展示,提供浏览式查询,可进行跨库检索。

(3) 超媒体检索(Hyper Media Retrieval)是对存储的文本、图像、声音等多种媒体信息的检索。它是多维存储结构,有向的链接,与超文本检索一样,可提供浏览式查询和跨库检索。

3. 按检索手段分,有手工检索和机器检索

(1) 手工检索(Hand Retrieval)是人直接用手、眼、脑组织、查找印刷型文献的检索,具有直观、灵活、无需各种设备和上机费用的优点。

(2) 机器检索(Machine Retrieval)又称计算机检索,是通过机器对已数字化的信息,按照设计好的程序进行查找和输出的过程。按机器检索的处理方式划分又有脱机检索和联机检索;按存储方式分,有光盘检索和网络检索。机检可大大提高检索效率,拓宽检索领域。

1.1.10 信息检索语言

1. 定义

检索语言是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包括:题名、责任者、号码、引文。

表述文献内容特征的语言包括:

(1) 分类语言。这指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 所示。

表 1-1 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
B	哲学
C	社会科学总论
D	政治、法律
E	军事
F	经济
G	文化、科学、教育、体育
H	语言、文字
I	文学
J	艺术
K	历史、地理
N	自然科学总论
O	数理科学和化学
P	天文学、地球科学
Q	生物科学
R	医药、卫生
S	农业科学
T	工业技术
U	交通运输
V	航空、航天
X	环境科学
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 1-2 所示为 T 大类的二级子类目），简表细分又构成详表，又称三级子类目。

表 1-2 T 大类的二级子类目

T	工业技术
TB	一般工业技术
TD	矿业工程
TE	石油、天然气工业
TF	冶金工业
TG	金属学与金属工艺
TH	机械、仪表工业

续表

TJ	武器工业
TK	能源与动力工程
TL	原子能技术
TM	电工技术
TN	无线电电子学、电信技术
TP	自动化技术、计算机技术
TQ	化学工业
TS	轻工业、手工业
TU	建筑科学
TV	水利工程

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

1) 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。

2) 关键词语言是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。

3) 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。包括人名、地名、书刊名等。

(4) 代号语言。包括专利号、标准号、合同号等。

1.1.11 信息检索的常用途径

信息检索的常用途径有：

(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。

(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。

(5) 顺序号途径。包括专利号、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。

1.1.12 信息检索的方法

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。

(1) 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。

(2) 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。

(3) 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。这也称为循环法。

1.2 计算机信息检索基础

1.2.1 计算机检索概述

1. 计算机检索的概念

通过计算机进行的文献信息检索称为计算机检索。

随着计算机技术、远程通信技术和信息存储技术的飞速发展,信息检索由手工检索过渡到计算机信息检索。计算机检索的成功应用,为我们更为及时、准确、全面地继承、利用和发展人类的科研成果提供了先进的手段。

2. 计算机检索发展

(1) 第一阶段:脱机检索阶段(20世纪50年代至60年代)。

脱机检索(Offline Retrieval)即批处理检索,检索有三个环节:检索要求、检索系统、检索结果。具体检索过程为检索人员→检索策略→成批检索→用户。

这个检索阶段的缺点是:

- 1) 地理上的障碍(远,不便于检索结果的获取)。
- 2) 时间上的迟滞(定期检索,不能及时获取)。
- 3) 封闭式的检索(指检索策略一经输入系统就不能更改,更不能依据机检应答来修