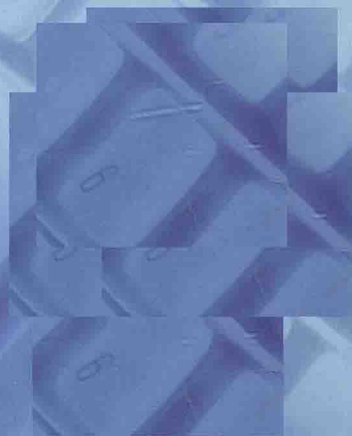


实用文献信息资源检索

闫军秀 主编



河南人民出版社

实用文献信息资源检索

闫军秀 主编

河南人民出版社

主 编：闫军秀

副主编：（按姓氏笔画为序）

刘惠萍、吴起立

编 委：（按姓氏笔画为序）

叶 国 闫军秀 刘惠萍

陈宝昀 吴起立 张长琳

图书在版编目（CIP）数据

实用文献信息资源检索 / 闫军秀主编. — 郑州：河南人民出版社，2006. 1

ISBN 7-215-04540-4

I. 实... II. 闫... III. 情报检索 IV. G252. 7

中国版本图书馆 CIP 数据核字（2006）第 007083 号

河南人民出版社出版发行

（地址：郑州市经五路 66 号 邮政编码：450002 电话：65723341）

新华书店经销 黄委会设计院印刷厂印刷

开本 787 毫米×1092 毫米 1/16 印张 12.875

字数 320 千字

2006 年 1 月第 1 版 2006 年 1 月第 1 次印刷

定价：26.00 元

前 言

21 世纪是信息的世纪,是知识经济的时代,信息对人们的工作、学习、研究乃至日常生活都产生着深远影响。社会信息化日益成为衡量一个国家、地区和民族综合实力的重要标志。当代社会的竞争也是信息的竞争,有效的信息是竞争取胜的关键因素,谁控制了信息,谁就掌握了先机。因此,在新的信息环境下,人们必须具备感知信息、获取信息、评价信息、利用信息的能力。如何培养和提高人们的综合信息素质,提高信息意识,掌握信息检索的方法和技巧,学会从海量的信息中获取自己所需信息的本领,增强自学能力和科研能力,就变得至关重要。为此我们编写了此书。

全书共分 7 章。第 1 章和第 2 章介绍了进行信息检索的基本理论知识;第 3 章介绍了国内主要的计算机检索系统;第 4 章和第 5 章分别介绍了国外综合性和专业性检索系统;第 6 章介绍了国外网络数据库检索系统;第 7 章介绍了特种文献检索系统。

本书由闫军秀任主编,刘惠萍、吴起立任副主编。各章节的执笔者分别为(按姓氏笔画顺序排):郑州大学图书馆叶国(第 2 章第 1 节、2 节和第 4 章),河南质量工程职业学院闫军秀(第 6 章和第 7 章第 1 节),偃师市图书馆刘惠萍(第 1 章和第 3 章第 1 节、2 节、3 节),信阳市图书馆陈宝昀(第 5 章第 2 节和第 7 章第 2 节、3 节),郑州轻工业学院图书馆吴起立(第 3 章第 5 节和第 5 章第 1 节、3 节),河南公安高等专科学校图书馆张长琳(第 2 章第 3 节和第 3 章第 4 节)。

在本书的编写过程中,参考和引用了近年来最新的文献信息研究成果、论著和论文及一些资料,书中采用的部分资料和图表取材于他们的研究成果和著作,有的在参考文献中列出,有的没有列出,我们在此向他(她)们表示深深的敬意和感谢。

由于编者水平有限,加之时间仓促,疏漏和不妥之处在所难免,诚望专家和读者批评指正,以便我们进一步的修改和完善。

编著者

2005 年 10 月

目 录

第1章 信息检索概论	1
1.1 信息、知识、情报、文献的基本概念	1
1.1.1 信息	1
1.1.2 知识	2
1.1.3 情报	2
1.1.4 文献	2
1.2 信息检索的目的和意义	4
1.3 信息检索基本知识	5
1.3.1 信息检索概念	5
1.3.2 信息检索语言(检索标识系统)	5
1.3.3 检索工具	9
1.3.4 检索方法	10
1.3.5 检索步骤(程序)	10
1.3.6 检索策略	11
1.3.7 参考工具	11
1.3.8 参考工具与检索工具的异同	12
1.4 计算机检索的基本知识	12
1.4.1 计算机检索的发展	12
1.4.2 计算机检索特点	12
1.4.3 计算机检索基本功能	13
1.4.4 计算机检索策略	14
第2章 Internet 信息检索	17
2.1 Internet 概述	17
2.1.1 Internet 的起源与发展	17
2.1.2 中国与 Internet	17
2.1.3 Internet 的接入方式	18
2.1.4 统一资源定位器(URL)	18
2.2 网络信息资源	18
2.2.1 万维网信息资源	19
2.2.2 FTP 信息资源	19
2.2.3 TELNET 信息资源	19
2.2.4 USENET 信息资源	19
2.3 国内外常用的 WWW 搜索引擎	20
2.3.1 雅虎中国	20
2.3.2 新浪搜索引擎	24
2.3.3 263 搜索引擎	29

2.3.4	网易搜索引擎	31
2.3.5	搜狐搜索引擎	33
2.3.6	Google 中文	35
2.3.7	Infoseek 搜索引擎	38
2.3.8	Excite 搜索引擎	39
2.3.9	Metacrawler (元搜索引擎)	44
第3章	国内主要计算机检索系统	48
3.1	CNKI 中国期刊网全文数据库	48
3.1.1	CNKI 数据库种类	48
3.1.2	CNKI 数据库使用条件	49
3.1.3	CNKI 检索方法	49
3.2	重庆维普中文科技期刊数据库	52
3.2.1	维普数据库种类	52
3.2.2	维普数据库使用条件	52
3.2.3	维普数据库使用方法	53
3.2.4	注意事项	55
3.3	万方数据资源系统	56
3.3.1	商务信息子系统	57
3.3.2	科技信息子系统	59
3.3.3	数字化期刊子系统	62
3.4	中国资讯行	64
3.4.1	专业数据库简介	65
3.4.2	使用方法	65
3.4.3	检索结果	68
3.5	中国科学引文数据库	69
3.5.1	CSCD 的检索方法	70
3.5.2	检索结果显示	74
3.5.3	记录采集	75
3.5.4	检索结果的输出和打印	75
3.5.5	检索策略管理	76
第4章	国外综合性检索系统	77
4.1	科学引文索引	77
4.1.1	概述	77
4.1.2	印刷版《科学引文索引》编排方法	77
4.1.3	《SCI》查找方法举例	80
4.1.4	《科学引文索引》光盘检索	81
4.2	美国《工程索引》	84
4.2.1	概述	84
4.2.2	印刷版《工程索引》检索方法	84
4.2.3	《工程索引》光盘检索	86
4.2.4	Ei Village 检索	90

4.3 英国《科学文摘》	95
4.3.1 概述	95
4.3.2 《科学文摘》手工检索方法	96
4.3.3 《科学文摘》网上数据库检索	99
第5章 国外专业性检索系统	107
5.1 《化学文摘》(CA)	107
5.1.1 概况	107
5.1.2 《化学文摘》手工检索	107
5.1.3 CA光盘数据库检索	118
5.2 《生物学文摘》(BA)	125
5.2.1 概述	125
5.2.2 《生物学文摘》手工检索	125
5.2.3 BIOSIS PREVIEW 数据库检索	130
5.3 《食品科技文摘》(FSTA)	136
5.3.1 FSTA 手工检索	136
5.3.2 FSTA 光盘检索	139
第6章 国外网络数据库检索系统	144
6.1 OCLC 和 First Search 数据库	144
6.1.1 First Search 的主题范畴和 14 个基本数据库简介	144
6.1.2 First Search 使用方法	146
6.2 UnCover 数据库检索	152
6.3 《剑桥科学文摘》(CSA) 数据库	156
第7章 特种文献检索系统	162
7.1 专利文献检索	162
7.1.1 专利及专利文献概述	162
7.1.2 中国专利信息检索系统	163
7.1.3 国外专利信息检索系统	170
7.2 学位论文及其检索	176
7.2.1 概述	176
7.2.2 中国高等学校学位论文检索信息系统	177
7.2.3 美国《国际学位论文摘要》手工检索	179
7.2.4 ProQuest Digital Dissertation (PQDD) 网络数据库检索	180
7.3 会议论文检索	185
7.3.1 概述	185
7.3.2 国内会议文献检索	186
7.3.3 国外会议文献检索	186
7.3.3.1 《World Meeting》手工检索	186
7.3.3.2 《Index to Scientific & Technical Proceedings》手工检索	188
7.3.3.3 《ISTP》光盘检索	192
主要参考文献	198

第 1 章 信息检索概论

1.1 信息、知识、情报、文献的基本概念

1.1.1 信息

信息一词在中国历史文献中最早见于唐诗，拉丁词源是 *informatio*，英文词义是 *information*；20 世纪中叶以后，其本质才不断被揭示，并被引入哲学、信息论、系统论、控制论、传播学、情报学、管理学、通信、计算机科学等领域。围绕信息定义所出现的流行说法不下百种，哲学家认为信息是物质的普遍属性，是认识世界的依据；数学家认为信息是使概率分布发生变动的东西；各类学者分别从不同角度定义为“信息是选择的自由度”、“信息是人们在适应外部世界并且使这种适应反作用于外部世界的过程中同外部世界进行交换的内容的名称”、“信息是事物之间的差异”、“信息就是消息、信号、数据”等等。总之作为日常用语，信息是指音信、消息；作为科学术语，广义指事物属性的表征，狭义指系统传输和处理的对象。

信息无处不在，无时不有，普遍存在于人类社会和自然界中，所以常把信息分为自然信息和人工信息。自然信息是体现自然界物质特征和现象的信息，包括生物信息和物理信息。人工信息是指人类在社会实践中产生、加工、发出的信息。它既包括体现人类活动和社会发展的信息，也包括人类接收并破译的自然信息。对于信息，无论是自然信息还是人工信息，它们都具有以下基本属性。

(1) 普遍性 信息广泛存在于自然界、人类社会及思维领域中。只要有事物存在，只要有事物的运动，就会有信息存在。

(2) 客观性 就世界整体而言，信息的根源是物质世界，信息的存储、传播依靠物质和能量。

(3) 中介性 就物质世界的层次来看，信息既区别于物质，又区别于精神。物质是信息的载体，物质的运动是信息的源泉。信息来源于物质，又不是物质本身，它可以脱离源物质而相对独立地存在；信息也来源于精神世界，但又不限于精神的领域。

(4) 无限性 无论是在无限还是有限的空间里，随着时间的无限推移，事物的发展变化是无限的，信息也将无限扩充。

(5) 扩散性 信息可以在时间上或空间上从一点转移到另一点，可以通过不同媒介进行传播。

(6) 可知性 信息不是虚无缥缈的东西，也不是可以随意想象和创造的事物。信息是可为人们所感知和感受并利用的。

(7) 可贮性 信息可以被收集、加工、整理、筛选、归纳、综合，并可通过记忆和各种载体来载荷。

(8) 共享性 信息可以多向多次传播，为人们所共享，其信息量不会因传播的过程而减少，也不会因分享者的增多而减少。

1.1.2 知识

知识是人类在认识和改造世界的社会实践中获得的对客观事物本质和运动规律的认识。在生活、生产和科研等活动中,人脑通过对客观事物发出的信息的接收、选择、处理,得到对事物一般特征的认识,形成了概念。在反复实践和认识的过程中,人脑通过对相关概念的判断、推理、综合,加深了对事物本质的认识,构成了人们头脑中的知识。

知识的类型可以从不同角度划分:按反映对象的深刻性可分为生活常识,科学知识;按成熟程度可分为经验知识,理论知识;按存在方式可分为主观知识,客观知识;按可用性可分为个人知识,社会知识;按门类结构可分为基础知识,技术知识,应用知识;按描述内容可分为哲学知识,自然科学知识,社会科学知识,思维科学知识等。知识的类型无论怎样划分,作为知识都具备以下基本属性。

(1) 实践性 社会实践是一切知识产生的基础和检验知识的标准,科学知识对实践有重大指导作用。

(2) 规律性 人们对实践的认识是一个无限的过程,人们获得的知识在一定层面上揭示了事物及其运动过程的规律性。

(3) 渗透性 随着知识门类增多,各种知识可以相互渗透,形成许多新的知识门类,形成科学知识的网状结构体系。

(4) 继承性 每一次新知识的产生,既是原有知识的深化与发展,又是更新的知识产生的基础和前提。知识被记录或被物化为劳动产品后,可以世代相传和利用。

1.1.3 情报

情报的概念在不同历史时期有着不同的涵义。人们最早认为情报是战时关于敌情的报告;19世纪到20世纪50年代认为情报是系列化的知识;20世纪后期至今认为情报是人们为一定目的搜集的有使用价值的知识或信息。也如著名科学家钱学森所说:“情报是活化了的知识”,即被利用的知识。情报具有以下基本属性。

(1) 知识性与信息性 情报必须具有实质内容,凡人们需要的各种知识或信息都可成为情报的内容。

(2) 针对性 任何知识与信息,人们不利用、不知道其存在,就不能成为情报。情报处于运动状态中,是针对特定用户、特定需求提供的可利用的知识与信息。

(3) 时效性 情报只有在合适的时间内才会产生更大效益。如果随着时间的推移,提供的情报失去了使用价值,这样的情报也只能称之为知识或信息了。

1.1.4 文献

国际标准化组织《文献情报术语国际标准》(ISO/DIS 5217)对文献的描述是:“在存贮、检索、利用或传递记录信息的过程中,可作为一个单元处理的,在载体内、载体上或依附载体而存贮有信息或数据的载体”。我国颁布的《中华人民共和国国家标准·文献著录总则》(GB 3792·1—83)对文献定义为:“文献是记录有知识和信息的一切载体”。

1.1.4.1 文献有三个基本要素构成

(1) 知识信息内容 这是文献的核心与灵魂。

(2) 信息符号 揭示和表达知识信息的标示符号,如文字、图形、数字、声频视频等。

(3) 载体材料 是记录知识信息符号的物质材料,如龟甲兽骨、纸张、胶片胶带、光盘磁盘等。

1.1.4.2 文献的分类

(1) 按文献的加工深度分类 有零次文献、一次文献、二次文献、三次文献。

- 零次文献：是指未经刊载或未经公开交流的最原始的文献，如私人笔记、实验记录、设计草图、论文草稿、书信等，是一次文献的素材。

- 一次文献：也称原始文献，是出自作者本人的研究成果，并在社会上公开交流或发表的文献。它是最主要的文献情报源，是产生二次、三次文献的基础。

- 二次文献：是指在一次文献的基础上，对分散无序的一次文献进行收集、整理、加工，把文献的外部特征和内容特征著录下来，压缩成有组织、有序化的文献，如书目、索引、文摘、题录等。它是用来查找一次文献的“检索性文献”，是重要的指示性情报源。

- 三次文献：在一次文献和二次文献的基础上，通过分析、综合、提炼、重组而形成的再生文献，如主题述评、综述研究、综述报告、百科全书、文献指南、数据手册等。它源于一次和二次文献，又高于一次和二次文献，是情报研究的产物和成果，是人们掌握情报源的主要资料。

(2) 按文献的载体形式分类 有印刷型、缩微型、机读型、声像型、数字化型文献。

- 印刷型：是以纸张为存贮介质，以印刷为记录手段的文献形式，是最常见的传统的文献形式。其优点是便于携带和阅读。缺点是体积大、长期保管困难。

- 缩微型：是以感光材料为载体，利用摄影等光学记录技术使印刷型文献缩小许多倍的文献形式。包括缩微胶片、缩微胶卷和缩微卡片等。其优点是再现度高、体积小、易保管、价格便宜等，缺点是保存条件高、阅读不方便。

- 声像型：是运用录音、录像和摄影技术直接记录声音与图像的文献形式，包括唱片、录音带、录像带等。其优点是图文声并茂。

- 机读型：是指将文字和图像转化成二进制数字代码，记录在磁带、磁盘等载体上的文献形式。阅读时，再用计算机输出，转换成文字或图像。其优点是易于处理，存取速度快。缺点是需要借助计算机，价格昂贵，保存条件要求高。

- 数字化型：又可分为光盘型、联机网络型。光盘型文献是采用数字化存储方式，用激光来记录和再现信息的高密度存储介质制成，可同时存储声音、图像和文字等。其优点是存储量大、体积小、重量轻且携带方便。联机网络型是指以电子方式或机读方式生产和发行的、并通过电子计算机输出设备和电信网在视频终端上显示出来的文献。这包括 20 世纪 70~80 年代建立起来的计算机联机检索系统，如美国的 DIALOG、欧共体的 ESA 和德国的 STN 等联机系统，也包括近十多年迅速发展起来的 Internet WWW 网络信息资源。

(3) 按文献的出版形式分类 有图书、连续出版物和特种文献。

图书包括专著、工具书、教科书、史书、古籍等；连续出版物包括期刊、报纸、学报等；特种文献包括专利文献、会议文献、标准文献、科技报告、政府出版物、产品样本等。

1.1.4.3 科技文献的特点

科技文献是人们从事生产斗争和科学实验的记录。它是人类共同的精神财富。17 世纪末，许多科学协会相继成立，促进了会员之间的学术交流活动，为了能在广泛的范围内了解和推广新的发明创造，于是就出现了科技杂志。由于工业生产的发展，各资本主义国家从 17 世纪后半期和 18 世纪开始相继成立了专利局，审理创造发明专利，于是出现了专利文献。到 19 世纪中叶，才陆续出版了由学会创办的会志。到 20 世纪 60 年代，由于科学技术的高速发展，科技文献也剧增，尤其是近 20 年来，文献数量和品种更是都达到了历史的最

高峰。现代科技文献总的特点有以下几方面。

- 文献数量增长迅猛：据统计，目前世界上每年产生的图书 80 万~90 万种，期刊 20 万种，研究报告 70 余万篇，会议论文 10 多万篇，政府出版物超过 20 万篇，标准文献 75 万件，专利文献 105 万件，产品样本 50 万件，技术档案数百万件，学位论文 10 万件。

我们可从美国《化学文献》(CA) 历年收录文摘条数的变化来了解文献数量增长的速度：

1907 年	7975 条	1917 年	15601 条	1927 年	32909 条
1937 年	63038 条	1947 年	38386 条	1957 年	101027 条
1967 年	243982 条	1977 年	409841 条	1987 年	477177 条
1997 年	716564 条	1998 年	681008 条	1999 年	720084 条
2000 年	725195 条				

- 文献种类繁多：在现代科技文献中，除传统印刷型文献外，缩微型、机读型、声像型、多媒体及光盘等新型载体发展十分迅速，已呈现文献数字化、电子化的大趋势。

- 文献分布分散：由于学科间的相互交叉、相互渗透，使许多专业文献发表在非专业期刊中。

- 文献的半衰期缩短，失效快：科技发展得越快，文献的新陈代谢就越快。据报道，科技文献的有效期已从 19 世纪的 50 年，缩短到如今的 3~8 年，使用时效大大缩短。尤其是报道新技术、新工艺、新材料的文献，半衰期更短，失效更快。

1.2 信息检索的目的和意义

随着科学技术不断地发展、进步、日新月异，科技信息、科技文献也随之快速增长，尤其是 Internet 的出现，网络信息更是如潮水汹涌，势不可挡。如何从浩如烟海的信息中迅速准确地获取自己所需的信息资料，就必须掌握信息检索的方法。信息检索方法是每个人打开知识宝库的金钥匙。一个人不管他有多么高明、多么具有研究和开发能力，都不可能孤立的存在、孤立的研究，他都必须在前人研究的基础上继承和借鉴，从而有所发明、有所创造。信息检索的目的和意义就在于：

(1) 启迪创造性思维 文献资料既是过去经验的总结，又是未来的向导，全世界一切有成就、有贡献的科学家，都是在广泛吸收他人的知识、得到启迪而取得的。

(2) 提高能力 自我知识更新能力、自学能力和研究能力。

(3) 避免重复劳动，节省科研开支，推动科技进步 任何一项科学研究都离不开信息检索、文献查阅，通过信息检索和文献查阅以了解和掌握国内外是否有人做过或正在做类似研究，采用什么方法，取得什么结果，有什么经验教训，尚存在什么问题，以便借鉴、改进和部署自己的研究方向和内容。只有这样才能做到心中有数，避免重复，少走弯路。并且可以把自己的研究工作建立在一个较高的起点上，缩短研究周期，获得较高水平的成果。

(4) 节省科研时间、加快科研步伐 掌握信息检索和文献检索的方法，可以节省研究人员查阅文献的时间。据统计，科研人员在他们从事科研活动中用于查找和翻阅文献信息的时间约占 1/3。如果熟悉检索方法，那么查阅文献的时间就会由 1/3 降至 1/5，反之，将会由 1/3 增至 1/2。因此，有效地掌握文献信息的检索方法，可以为科研工作赢得时间，加快科研速度，早出科研成果。

(5) 促进经济发展 在信息时代，信息的开发与利用已成为一个国家综合国力的重要标

志。在市场经济体系中,在中国加入 WTO 以后,竞争是通行的法规。任何个人、企业、单位和个人,要想在竞争中站住脚,都必须充分掌握情报资料,并利用它进行科学决策,这样才能在竞争中立于不败之地。

1.3 信息检索基本知识

1.3.1 信息检索概念

信息检索的定义:是从大量的信息集合中查找出所需信息的活动、过程与方法。这是一种针对广大信息用户来说的狭义说法。从广义上来说,信息检索包括信息的存储和检索两方面,即包括将信息按一定的方式组织存储起来和按信息用户需求查找出相关信息两方面。信息检索的本质是信息用户将自身的需求和信息集合进行比较和选择的过程。对于手工检索来讲就是将检索提问标识与检索工具书中的标引标识进行人工比较,选出相匹配信息的过程;对于计算机检索来讲就是将检索提问标识与存储在数据库中的标引标识自动进行比较,含有该标识相匹配的信息就会从数据库中并通过计算机输出的过程。

1.3.2 信息检索语言(检索标识系统)

1.3.2.1 检索语言的概念

语言是人类为交往而使用的一种符号体系。它以语音为物质外壳、以词汇为组合材料、以语法为组织原则。语言又可分为自然语言和人工语言。

自然语言是指人类在交际中自然演变而形成的语言,是约定俗成的语言,包括方言等。与“人工语言”相对。

人工语言是指人类为了特定目的,根据预先制定的一套规则编制而成的语言。其规则是预先明确规定的,包括计算机语言、信息检索语言、世界语等。与“自然语言”相对。

信息检索语言是根据检索需要而创制的人工语言,也称检索标识系统。它专门用于各种手工和计算机信息检索系统,其实质是表达一系列概括文献内容的概念及其相互关系的概念标识系统。也就是说,信息的存储和检索是两个紧密联系的过程,在存储和检索过程中要涉及到文献著者、标引者、检索者和情报用户四个方面的人。这四方面的人对同一文献的使用目的不同。为了使信息检索过程能顺利进行,在信息存储时,对信息资料的外部 and 内部特征要按照一定的语言来加以描述,而信息检索时也必须使用与其相同的语言来加以表达。这种把信息存储与检索联系起来、把标引人员与检索人员联系起来以便取得共同理解的语言,就叫检索语言或叫检索标识系统。

1.3.2.2 检索语言的类型

按一定标准对检索语言整体划分后形成的不同检索语言群体。各种检索语言的基本原理是一致的,只是在表达各种概念及其相互关系时所采用的方法不同,才形成了不同类型的检索语言。检索语言从不同角度划分种类很多,常用的有以下几种。

(1) 按表述文献特征划分 分为表述文献外表特征的检索语言和表述文献内容特征的检索语言两种。表述文献外部特征的检索语言主要包括篇名(书名)、著者、文献代码、引文等。表述文献内容特征的检索语言主要包括分类语言、标题词语言、关键词语言、叙词语言等。

(2) 按标识的组配方式划分 分为先组式检索语言、后组式检索语言和先组散组式检索语言。

① 先组式检索语言 是指检索标识在编表之前表述文献主题概念的标识已经固定组合

好的检索语言。如标题词语言、体系分类语言等。这种语言检索者比较习惯，适用于传统的检索工具。

② 后组式检索语言 是指检索标识在编表时没有预先固定组配，而是在检索时，根据检索的实际需要，按照组配规则临时进行组配的检索语言。如叙词语言、单元词语言等。这种检索语言适用于计算机检索系统。

③ 先组散组式检索语言 是指检索标识在编表时没有预先固定组配，而是在标引时组合成固定标识串的检索语言，如《美国国立医学图书馆医学主题表》。先组散组式检索语言的性能与先组式检索语言的性能相似。

(3) 按构成原理划分 分为分类检索语言、主题检索语言、代码检索语言和引文检索语言。

① 分类检索语言 是指用分类号表达各种概念，并将各种概念以学科性质为主加以划分和系统排列的检索语言。按编制方式可分为等级体系分类语言、组配分类语言和混合分类语言。

- 等级体系分类语言：它按学科体系的层次，从上到下，从总到分，逐级展开，各级类目预先固定组配，具有等级制结构。

- 组配分类语言：它用科技术语进行组配的方式来描述文献内容。这些科技术语按学科性质分为若干组，即“组面”。组面内各个术语都附有相应的号码。标引文献时，根据文献内容选择相应的组面和有关术语，把这些术语的号码组配起来，构成表达这一文献内容的分类号。

- 混合分类语言：它是将体系分类和组配分类相结合的一种检索语言。

在此重点介绍体系分类语言。

体系分类也称分类法或分类表，是使用历史最长的图书加工整理方法。我国常见的体系分类语言有《中国图书馆图书分类法》（简称中图法），《中国科学图书分类法》（简称科图法），《中国资料分类法》（简称资料法）。

下面主要介绍《中图法》分类体系。《中图法》将人类知识分为五大部分，22 个大类：

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	数学科学和化学	
P	天文学、地球科学	
R	药学、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、劳动保护科学	
Z	综合性图书	

} 社会科学部分

} 自然科学部分

在 T-工业技术大类下的 TP 类展开的下位类如下:

T	工业技术	TP 393	计算机网络
TP	自动化技术、计算机技术	TP 393.0	一般性问题
TP 3	计算技术、计算机技术	TP 393.01	计算机网络理论
TP 39	计算机应用	TP 393.02	计算机网络结构与计

从上表看出分类号的变化体现了各学科体系的从属关系, 分类号位数每增加一位, 则分类级别便低一级。如 TP 39 代表计算机应用, TP 393 类目是计算机网络, 是 TP 39 的下位类。确定分类号的方法有以下两种。

a. 直接查找法。熟悉分类表, 按学科体系从大类开始逐级查找。

- 对于只存在一个主题的课题, 方法就较简单, 如“机械手”。当要求的主题过于狭窄, 没有相应类目存在, 可归属在其上位类。

- 对于包含有两个以上主题内容的课题, 如“机械手的控制系统”涉及“机械手”和“控制系统”两个主题概念, 这时应考虑它们的主从关系, 在这个课题中研究的是仅针对机械手而言的控制系统, 可以说是控制系统在机械手中的应用, 当研究某种理论、方法、工艺等在某领域的应用时, 应分在其应用的类目, 即“机械手”所在类目。

- 对有两个主题, 又没有主次或应用关系的课题, 则在检索相关文献时应分别在两个分类号或两个主题领域中查找。

b. 间接获取分类号法。当手头有现成的对口文献或通过其他方法如主题方法查到的对口文献, 则可利用该文献标出的分类号, 进行分类查找。

② 主题检索语言 是指采用描述文献主题的语词标识并按字顺序列排检的检索语言。按主题词性质不同, 又可分为关键词语言、标题词语言、单元词语言和叙词语言。

a. 关键词语言。它是以关键词作为文献内容标识和检索的主题语言。关键词是指从文献中直接选取、未经规范处理的可表达文献主题内容的具有实际检索意义的语词。

目前, 关键词语言因不经过人为的筛选、能及时反应新科技用词, 所以较容易被检索者接受并喜欢, 且由于关键词是直接文献题目或正文中抽取的这一特征而被普遍应用于计算机检索中。但是人类语言中大量存在的一词多义、一义多词和同义词、近义词、相关词, 使关键词检索语言造成检索效率降低, 所以使用关键词检索时对其同义词、近义词和相关线索的收集较为重要。查找同义词、近义词和相关信息可采用下列几种方法。

- 定义查找法: 给一个词下定义, 实际上是用别的词组成句子去解释这个词, 被定义的词在意义上必定与定义它的词相关或相近, 所以定义的过程也是一个寻找同义词、相关词的过程。如, 抵触——格格不入——与另一方有矛盾。

- 语篇查找法: 这也是一个查找同义词、相关词的有效方法。一篇文章本身会聚集大量的同义词、相关词, 通过对一篇相关文章上、下文的分析, 会找出众多的同义词、近义词、相关词。

- 同构查找法: 是从词语表达的内容出发查找其相关词, 如“绿色包装”, 其中“绿色”隐含“环保”、“常温降解”、“易回收”等属性, 它们均是“绿色包装”的相关词、同义词。

b. 标题词语言。它是以标题词作为文献内容标识和检索的主题语言。标题词是指能直接表达文献主题和检索需求的、经规范化处理的语词, 是主题词的一种。

c. 单元词语言。它是以单元词作为文献内容标识和检索的主题语言。单元词又称元词。是指能表达文献主题的、经规范化处理的最小最基本的词汇单位, 其概念不可再分的词, 是

主题词的一种。

d. 叙词语言。它是以叙词作为文献内容标识和检索的主题语言。叙词又称描述词或叙述词。是以概念为基础，经过优选和规范化处理并具有概念组配和词间语义关系显示功能，用以表达文献主题和检索需求的语词，也是主题词的一种。

大多数印刷型检索工具均使用叙词语言作为主题标引，具体的表现形式即是叙词表。通常使用的叙词表有：国内大多检索工具使用的《汉语主题词表》、英国《科学文摘》使用的《INSPEC Thesaurus》、美国《工程索引》1993年后使用的《Ei Thesaurus》等。下面以英国《科学文献》使用的《INSPEC Thesaurus》为例介绍叙词表的著录形式。

《INSPEC Thesaurus》主要由两部分组成。

• 叙词主表：按叙词的英文字顺排列，每个叙词下都列出该词的使用范围说明、使用时间、上位词、下位词等相关信息，检索者根据从叙词表中获得的主题词信息，可以相应扩大或缩小检索用词范围，达到最佳检索效果。叙词主表编排格式如下：

Laser beam machining

UF Laser drilling

Laser machining

Laser trimming

BT Laser beam applications

Machining

TT Laster bean applications

Manufacturing processes

RT Laser beam etching

Laser materials processing

Micromachining

CC B0170G; B2575F; B4360B; B8620; C3355C

DT July 1973

PT Laser beam applications

Machining

各标头缩写含义如下表：

UF	使用范围，即 “Used for”	CC	分类号，classification code (s)
BT	上位词，Broader Term (s)	DT	引入日期，Date of Input
TT	族首词，Top Term (s)	PT	引入前使用的词，Prior Term (s)
RT	相关词，Related Term (s)	NT	下位词，Narrower Term (s)

• 词族表：此表反映了所有叙词之间的从属关系，以主表中的族首词字顺排列，前面小圆点越多，代表该词级别越低，任何一级带圆点的叙词是离它最近的、少一个圆点的上一级叙词的下位词。词族表编排格式如下：

laser beam applications

• Laser applications in medicine

• Laser beam machining

• Laser beam welding

- Laser materials processing
- • Laser ablation
- • • pulsed laser deposition
- • Laser beam annealing
-
- Plasma heating by laser
- Plasma production by laser
- • Laser fusion
- Remote sensing by laser beam

在检索课题时,选择规范化主题词应注意下列事项。

- 选事物名称或过程名称作为主题词:例如“蘑菇保鲜技术”,其中“蘑菇”是事物的名称,“保鲜”则是一种处理过程,这两个词均适合选择作为主题词。
- 避免使用一些泛指性词:如上例中的“技术”一词即属于泛指词,没有任何专指度,不合适作主题词。
- 适度专指性:有些过分偏窄的词也不适合作为主题词。例如上例中“蘑菇”一词可能在主题标引时因专指性太强,并未专门列出,这时应考虑它的上位类“食用菌”、“蔬菜”或“食品”等。

③ 代码检索语言 是用来标引、检索特定专业文献的某种代码系统。如化学物质登记号、专利号等,检索系统即是用代码检索语言来标引的。

④ 引文检索语言 是基于文献之间引证关系而形成的一种检索语言。它以引文为检索标识,根据引证关系将有关文献自然地偶合在一起。检索时通过引文标识可以回溯到一系列内容相关的文献。

1.3.3 检索工具

1.3.3.1 检索工具概念

检索工具是用以报道、存储和查找文献信息的一切工具与设备。它以各种类型的原始文献为素材,在广泛收集、筛选、分析的基础上,用选定的检索语言进行描述和标引,并按特定规则组织编排而成的二次文献。

作为检索工具都具有存储与检索两个方面的职能。一方面把有关文献的特征著录下来,成为一条条文献线索,并将它们进行系统排列,这就是所谓文献的存储过程。另一方面,检索工具能够提供一定的检索手段,使检索者可按照一定的检索方法,随时从中查找出所需要的文献线索,这就是文献的检索过程。

1.3.3.2 检索工具的类型

检索工具是一个综合的概念,种类很多,可从多角度进行划分。

(1) 按检索手段划分 分为手工检索工具、机械检索工具、计算机检索工具(即计算机检索系统)。

(2) 按载体形态划分 分为书本式检索工具、卡片式检索工具、缩微式检索工具、磁性材料式检索工具,如光盘、磁盘等。

(3) 按收录范围划分 分为综合性检索工具、专业性检索工具、专题性检索工具、单一性检索工具。

(4) 按著录对象和方式划分 分为目录、题录、文摘、索引、指南等。

(5) 按出版形式划分 分为期刊式检索工具、单卷式检索工具、附录式检索工具。

1.3.4 检索方法

针对某检索目的、为实现该检索计划或方案所采取的具体操作方法称之为检索方法。在检索过程中,应根据检索系统的性能、用户的实际需要和检索工作中遇到的实际情况,灵活应用各种检索方法。

在计算机检索条件下,常用的检索方法有积木法、引文珠串增长法、逐次分割法等。由于计算机检索速度快,所以在检索时,检索策略比检索方法显得更为重要(见1.4 计算机检索的基本知识)。在手工检索条件下,常用的检索方法有工具法、追溯法、循环法。

1.3.4.1 工具法

直接利用检索工具进行文献检索的方法,又称“常用法”。它包括顺查法、倒查法和抽查法。

(1) 顺查法 以课题研究的起始年代为起点,由远及近,利用检索工具逐年进行查找,一直查到最近期为止。利用该检索方法检索文献系统性较强,查全率较高,但费时费力。

(2) 倒查法 又称“逆查法”。即从近期向早期回溯,由近及远,逐年查找文献资料的方法。这种方法重点是获取近期文献或最新信息,查找时不需要一年一年地查到底,只需查到能满足基本需要的文献信息即可。该法优点是省时省力,但漏检率较高。

(3) 抽查法 针对研究课题处于兴旺时期的一段时间的文献信息进行查找。因为研究课题处于兴旺时期的研究成果最多,发表与报道的文献信息数量也最多。因此该法是一种付出少,检索效率和检索效果都较好的方法。

1.3.4.2 追溯法

利用已知文献的引用文献或参考文献为线索,由近及远,进行逐一追踪的查找方法,又称“追踪法”。该法查找方法简单,但漏检率较大,检索效率也不高。

1.3.4.3 循环法

将追溯法与工具法结合起来检索文献信息的方法。即利用检索工具查到某些文献,又利用这些文献所附参考文献追溯查找,这样分期分段地交替进行,循环下去,直到满足检索要求为止。采用此法检索效率较高。

1.3.5 检索步骤(程序)

文献检索一般可分以下五个步骤:即分析研究课题、选定检索工具、确定检索方案、进行具体查找、提供文献线索查阅原始文献。

(1) 分析研究课题 对检索的课题要进行周密的分析,弄清楚课题研究的目的、学科性质、主题内容、要达到的研究水平等,有的放矢地进行后续检索工作。

(2) 选定检索工具 通过分析研究课题,明确课题的检索范围和要求后,就要根据已确定的检索范围和要求进行检索工具的选定。当然要选择匹配的检索工具,这取决于检索者对检索工具熟悉和了解的程度。

(3) 确定检索方案 包括确定检索语言(标识)、检索途径、检索方法等。选定了检索工具后,就要选择与课题检索要求相匹配的检索标识,如分类号、关键词、标题词等。根据检索标识来选择检索工具提供的相应的检索途径。比如,课题要求检索的是专指性较强的文献信息,则最好选用主题标识的主题途径;若事先已知道本课题的研究专家,则可以利用著者途径。当然如果能将多种途径配合使用,将会收到更好的检索效果。确定检索途径后,还要根据课题的研究背景,确定检索方法,以提高检索效率。