

# 计算机 信息检索

李凌杰 冯有利 主编



中国物资出版社

# 计算机信息检索

李凌杰 冯有利 主编

中国物资出版社

**图书在版编目(CIP)数据**

计算机信息检索/李凌杰 冯有利主编. —北京:中国物资出版社,1999.6

ISBN 7-5047-1241-8

I. 计… II. 李… III. 计算机—信息检索 IV. G354

中国版本图书馆 CIP 数据核字(98)第 16379 号

中国物资出版社出版

北京市西城区月坛北街 25 号 邮编:100834

全国各地新华书店经销

郑州市中兴印刷厂印刷

开本:787mm × 1092mm 1/16 印张:12.25 字数:298 千字

1999 年 6 月第 1 版 1999 年 6 月第 1 次印刷

书号:ISBN 7-5047-1241-8/TP·0038

印数:0001-1000 册

定价:14.00 元

# 《计算机信息检索》

## 编 委 会

主 编 李凌杰 冯有利

副主编 谷 琦 杜建华 薛星梅 张运良 曾 仲

编 委 (以姓氏笔画为序)

王广三 冯有利 杰凌杰 李振龙

杜建华 谷 琦 张春梅 张维玲

张运良 郝小勤 曾 仲 翟继红

曹 杨 薛星梅

## 前 言

在即将跨入 21 世纪的今天,信息技术日新月异,信息高速公路以前所未有的速度在发展,如何从浩如烟海的信息海洋中及时地获取符合要求的、准确和全面的信息,是摆在信息工作者面前急需解决的问题。传统的文献检索教材已不能适应时代的发展和需要,为了更好地培养大学生和科技工作者的信息意识,提高他们利用计算机进行信息检索的能力,特邀请多年从事计算机信息检索课教学工作的教师编写此书。

本书系统阐述了信息检索的基本原理、信息检索工具;介绍了国外科技信息检索、专利、标准和科技报告检索,计算机信息检索技术、联机信息检索、光盘信息检索、计算机书目检索。

本书共分七章。各章节执笔者为:第一章:张维玲、张春梅、谷琦;第二章:张运良、杜建华;第三章:薛星梅、郝小勤;第四章:李振龙、曾仲、张运良;第五章:杜建华、谷琦;第六章:王广三、翟继红、郝小勤;第七章:曹杨、曾仲、薛星梅。

全部稿件经谷琦、杜建华、薛星梅、张运良、曾仲分别修改后,由李凌杰、冯有利统改定稿。

在编写过程中,参考了近年来最新的计算机信息检索教材和信息工作者的研究成果,在此向他们表示诚挚的谢意。该书从选题到出版过程中,自始至终得到了姜效先、夏英华、刘二灿、杜明、王均林、陈有富的大力支持,在此表示衷心地感谢。

本书虽然经著者努力,但错误之处在所难免,希望广大读者批评指正。

# 目 录

|       |                              |    |
|-------|------------------------------|----|
| 第一章   | 导论                           | 1  |
| 1.1   | 信息检索的概念                      | 1  |
| 1.2   | 信息检索的原理                      | 1  |
| 1.3   | 信息检索的类型                      | 2  |
| 1.4   | 信息检索的方法                      | 6  |
| 1.5   | 信息检索的步骤                      | 7  |
| 1.6   | 信息检索语言                       | 8  |
| 1.7   | 信息检索工具                       | 19 |
| 第二章   | 国外科技信息检索                     | 31 |
| 2.1   | 美国《工程索引》                     | 31 |
| 2.2   | 英国《科学文摘》                     | 39 |
| 2.3   | 美国《环境文摘》                     | 45 |
| 2.4   | 日本《科学技术文献速报》                 | 50 |
| 第三章   | 专利、标准及科技报告信息检索               | 57 |
| 3.1   | 专利文献及其检索                     | 57 |
| 3.1.1 | 专利文献概述                       | 57 |
| 3.1.2 | 专利文献分类                       | 61 |
| 3.1.3 | 中国专利文献及其检索                   | 63 |
| 3.1.4 | 国外专利文献检索工具<br>——《世界专利索引》及其检索 | 66 |
| 3.2   | 标准文献及其检索                     | 74 |
| 3.2.1 | 标准文献概述                       | 74 |
| 3.2.2 | 中国标准文献                       | 75 |
| 3.2.3 | 国际标准文献                       | 78 |
| 3.2.4 | 美国标准                         | 81 |

|                                |            |
|--------------------------------|------------|
| 3.2.5 英国标准 .....               | 83         |
| 3.2.6 日本工业标准 .....             | 84         |
| 3.2.7 法国标准 .....               | 85         |
| 3.3 科技报告及其检索方法 .....           | 86         |
| 3.3.1 科技报告概况 .....             | 86         |
| 3.3.2 美国的科技报告 .....            | 87         |
| 3.3.3 美国科技报告的检索 .....          | 89         |
| 3.3.4 中国的科技报告 .....            | 90         |
| 3.3.5 国内外其他科技报告检索工具要目 .....    | 91         |
| <b>第四章 计算机信息检索技术 .....</b>     | <b>92</b>  |
| 4.1 概 述 .....                  | 92         |
| 4.1.1 计算机信息检索系统的逻辑构成 .....     | 94         |
| 4.1.2 计算机信息检索系统的物理构成 .....     | 96         |
| 4.1.3 信息检索的数学模型 .....          | 98         |
| 4.2 文献数据库 .....                | 105        |
| 4.2.1 书目数据库 .....              | 106        |
| 4.2.2 全文数据库 .....              | 108        |
| 4.3 文献数据库的建设与维护 .....          | 110        |
| 4.4 非文献型数据库 .....              | 114        |
| 4.4.1 数值数据库 .....              | 114        |
| 4.4.2 指南数据库 .....              | 115        |
| 4.4.3 术语数据库 .....              | 115        |
| 4.4.4 图像数据库 .....              | 117        |
| 4.4.5 全息数据库 .....              | 118        |
| 4.5 计算机信息检索策略与步骤 .....         | 120        |
| <b>第五章 联机信息检索 .....</b>        | <b>125</b> |
| 5.1 国内联机信息检索 .....             | 125        |
| 5.1.1 中国科技信息所的 ISTIC 系统 .....  | 125        |
| 5.1.2 北京文献服务处的 BDSIRS 系统 ..... | 125        |
| 5.1.3 化工部的 CHOICE 系统 .....     | 126        |
| 5.1.4 机电部的 MEIRS 系统 .....      | 127        |

|             |                          |            |
|-------------|--------------------------|------------|
| 5.2         | 国际联机信息检索 .....           | 128        |
| 5.2.1       | 国际联机信息检索原理及服务方式 .....    | 128        |
| 5.2.2       | 几种国际联机检索系统简介 .....       | 129        |
| 5.2.3       | 因特网信息检索 .....            | 129        |
| <b>第六章</b>  | <b>光盘信息检索 .....</b>      | <b>144</b> |
| 6.1         | 光盘检索概述 .....             | 144        |
| 6.1.1       | 光盘的基本原理 .....            | 144        |
| 6.1.2       | 光盘的种类 .....              | 144        |
| 6.1.3       | 光盘数据库检索方法 .....          | 145        |
| 6.1.4       | 光盘检索系统的构成 .....          | 146        |
| 6.2         | 美国《工程索引》光盘数据库检索 .....    | 147        |
| 6.3         | INSPEC 光盘检索 .....        | 153        |
| 6.4         | 中文科技期刊篇名数据库检索 .....      | 158        |
| 6.5         | 中国学术期刊光盘检索 .....         | 160        |
| 6.6         | 中国专利文献光盘数据库检索 .....      | 164        |
| 6.7         | 万方数据库检索 .....            | 167        |
| <b>第七章</b>  | <b>计算机书目检索 .....</b>     | <b>174</b> |
| 7.1         | 计算机书目检索概述 .....          | 174        |
| 7.2         | OCLC 信息检索 .....          | 175        |
| 7.3         | 北京图书馆“文津”文献管理系统 .....    | 178        |
| 7.4         | 深圳图书馆自动化集成系统(ILAS) ..... | 181        |
| 7.5         | 丹诚图书馆集成系统 .....          | 184        |
| 7.6         | 博菲特文献管理集成系统 .....        | 186        |
| <b>参考文献</b> | <b>.....</b>             | <b>190</b> |

# 第一章 导 论

## 1.1 信息检索的概念

信息检索就是从众多信息中查找并获取信息的过程。从计算机系统来说,信息检索就是用户根据某种目的,在一定的时间内,利用计算机从已整理好的存储信息——数据库中,得到所必需的信息。早期的信息检索,基本上是一种对存储信息的查找。但是进入 90 年代以后,它正从低级阶段向高级阶段发展,即:“智能检索”阶段。它不仅能对已有存储信息实现高速查找,而且还可以进行类比、组合、联想、推断、创造等高级活动。

信息检索有两种方法。一是直接从信息源和信息文献载体中获取信息,称为直接检索;二是通过信息检索工具的指引获取所需的信息,称为间接检索。

直接检索是信息用户常用的检索方法。用这种方法检索时,可以明确判定所检索到的信息是否符合需要,且对于非文献载体的信息检索来说比较迅速、方便。但对于现代大规模的信息检索,很难广、快、精、准地查到所需的全部信息资料。

间接检索克服了直接检索的缺点。它把信息资料“贴上”分类号或主题词等多种检索标志,并按照这些标志把信息有效地组织起来,在信息检索时利用检索标记可以查到所需信息的线索或信息本身。

在信息检索时,片面强调间接检索而忽视直接检索或强调直接检索而忽视间接检索的做法是不正确的。科学的方法是根据信息本身的传播渠道,把直接检索和间接检索有机地结合起来。

## 1.2 信息检索的原理

直接检索的原理比较简单,即通过一定方法(如情报调研等)从信息源收集信息或直接浏览信息的文献载体获取信息。

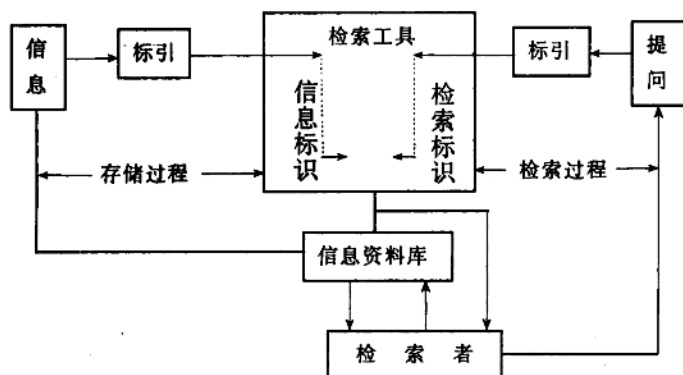


图 1—1 信息检索系统原理图

间接检索要通过信息检索系统来实现。间接检索的原理如图 1-1 所示。信息检索者的检索提问标识与存储在信息检索工具中的信息特征标识进行相符性比较。若检索标识与信息标识相符,则具有该特征标识的信息就从检索工具中输出。输出的信息就是初步命中检索所需的信息。

从图 1-1 不难看出,信息检索系统包括存储系统和检索系统两大部分。信息通过存储系统使用户所需信息从检索工具中输出。显然,存储过程是实现检索过程的前提和基础,而检索过程实现了存储过程的目的和意义,两者相辅相成,互为依存关系。

信息检索系统运作的全过程包括存储和检索两个分过程。存储过程就是运用检索语言对信息的外表特征和内容特征进行标引,形成信息的特征标识,按照存储所提供的检索途径,查获与检索提问标识相符合的文献特征标识的过程。显然,检索过程是存储过程的逆过程。

## 1.3 信息检索系统的类型

### 1.3.1 信息需求

前面已经提出,信息的载体可分为文献型和非文献两大类。基于此,我们可以把用户的需求归纳为图 1-2 所示结构。

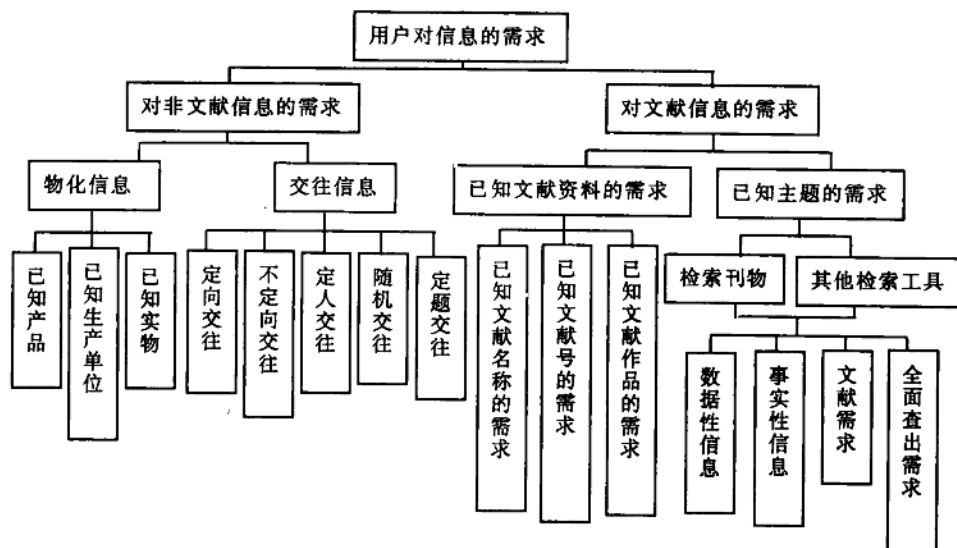


图 1-2 用户信息需求类型

用户的信息需求按其内容可分为 3 种类型:

(1)知识型信息需求。这类需求要解决“是什么”、“怎么做”的问题,如计算机知识、管理知识等。这类信息的需求可以增加需求者的知识或增长其才干。

(2)消息型信息需求。如经济市场信息、科技信息等。它是一种动态信息,可作为用户在决策时的参考。

(3)数据、事实与资料型信息需求。如自然常数、统计数据、组织机构情况、某一事件的记载等,为信息用户查找某一数据、事实与资料而用。

### 1.3.2 信息检索系统的类型

信息检索系统是由一定的设备和信息集合构成,面向一定的用户,具有信息采集、组织、存储、选择和传播等功能的信息服务设施。

根据不同的标准,信息检索系统可以划分为许多不同的类型。

#### 1. 按设备划分

根据信息存储与检索使用的设备,它可分为以下类型:

##### (1)书本式检索系统

它包括一切以传统书刊形式提供为人们查找各种信息或数据的出版物,如文摘杂志、题录或索引刊物、书本式目录、参考工具书等。它们是信息检索计算机化以前的信息检索系统的主要形式,而后又成为机检系统数据库的基础。目前,这类系统在信息检索中仍发挥着重要作用。

##### (2)卡片式检索系统

它包括一切以普通卡片存储和查找信息的工具,如图书馆内的各种卡片目录,管理部门的各种卡片档案。与书本式系统相比,它较便于信息累积和更新,适用于单位或个人自建自用,成本较低,在信息检索和资料管理中曾发挥过重要作用。目前,它已受到计算机化目录系统(如联机共用目录查询系统 OPAC 和数据库管理系统)的有力挑战,面临着被取代的命运。

##### (3)穿孔卡片检索系统

穿孔卡片(punched card)是一种长方形卡片,在边缘或中央部分穿孔表示和记录信息,借助手工或机器对穿孔卡片集合进行分拣(检索)。它是信息检索机械化的早期产物,出现于本世纪 40-50 年代,后被计算机检索系统取代。

##### (4)缩微式检索系统

它是以缩微照相制品(胶卷、胶片)为信息存储介质的检索系统。缩微胶卷检索系统分手工、半自动和自动式 3 种。手工检索又分靶标方式、指示线方式和顺序编码方式。靶标(又称闪现卡片)是拍摄时设立的检索标志,标明缩微画面的组号,便于读者选择所需画面。指示线也是用来标示画面地址的编码方式。检索时,使胶卷在阅读器屏幕上快速通过,当指示线与刻度尺上某一预定刻度相一致时,即可检出所需画面。顺序编码方式就是拍摄时先给每幅画面顺序编号,借助索引查出特定画面的顺序号,然后检出相应的画面。半自动检索方式按画面距胶卷传送长度的计数器,当计数器上显示的长度与索引中查到的距离数字相符时,阅读器停止转动,屏幕上显示出所需的画面。故它又称为长度表或量程表检索法。自动检索方式包括光点法和二进制编码法。光点是拍摄时在画面下方设置的方形标记。检索前,先在控制键盘上输入所要查的资料的光点顺序数,然后使胶卷转动起来,当阅读器上的光电计数器记录挑点数与输入数字一致时,输卷装置自动停转,所需资料就会显示在屏幕上。二进制编码法用黑白方块的组合表示二进制代码,拍摄时自动编码,表示画面的地址,检索方法与光点法相似,但需使用专门的缩微摄影机和检索阅读设备。

缩微平片检索也有手工、半自动和自动式三种。手工方式有标题索引方式、颜色分类方

式和槽口索引方式。半自动检索几乎都是采用在平片边缘开槽口,利用磁铁吸力选出所需平片。自动检索有齿槽夹具和固定位置两种,检索机构与阅读器连接,可自动检出所需平片并自动显示在阅读器屏幕上。近来,计算机辅助缩微品检索系统逐步普及,已出现计算机与缩微平片阅读器自动连接的系统,可自动检出和显示缩微画面,最后获得硬拷贝。

#### (5) 计算机信息检索系统

早期的计算机信息检索系统以机器穿孔卡片为存储介质。后来,发展到以磁带为外存储器,利用计算机顺序读取磁带数据库中的数据,故又称顺排档检索系统。60年代中期以后,随着磁盘和大容量存储器投入使用,大量数据可以随机存放在磁盘上,基于磁盘的检索系统正式出现。它可以支持回溯检索。人们又称之为倒排档检索系统。计算机检索系统的高级形式是联机检索系统。它通常由大型计算机系统、数据库和一批检索终端等设备构成。检索终端通过通讯线路与系统的主机连接,并在中央处理机控制下工作,查询系统的数据库。

#### (6) 光盘检索系统

它出现于80年代,以大容量光盘存储器为数据库的存储介质,利用光盘驱动器和微机读取和检索光盘上的数据。

#### (7) 电视信息查询系统

它是数字通讯、电视和电子计算机3种技术相结合的产物,分为广播式和交互式两种。广播式电视信息查询系统(teletex或teletext)利用电视广播网循环传送专门的信息页(即记录有用户需要的数据的电视画面)。电视用户用一个简单的按钮选择装置来检索所需的数据画面。可查询的信息内容通常有天气预报、体育新闻等。国内常把这种系统称为图文电视和用户电视报文。

交互式电视信息查询系统(videltex;videodata)用电话线或电缆连接计算机和用户的电视接收机,是一种双向交互式 and 闭合型的系统,与联机检索系统相似。数据库中存储的信息也是以页(frame,帧)为单位,每页有编号。用户通过与系统对话,查询系统的数据库。检索方式多为菜单选择式,也可以按页号或关键词来检索。交互式系统的应用范围很广,如电话号码查询、订票、电子邮件、电子购物、银行资金转帐、旅游服务、教育培训、娱乐和广告等。

#### (8) 多媒体检索系统

指建立在多媒体数据库(multimedia database)基础上的情报检索系统。

### 2. 按功能划分

各种信息或数据查询都属于信息检索的范畴。许多信息系统都含有检索的机制和过程“它们存储的数据文件或信息库,在响应用户提问时被处理,然后生成和输出答案。在这个意义上说,它们与普通的信息检索系统是相同的。我们可以把它们看成是广义的检索系统。虽然它们的内部逻辑结构和功能特征有差异,但相互之间联系紧密,且有相互融合的趋势。

#### (1) 文献检索系统

文献检索系统(document retrieval system)是狭义的信息检索系统。目前绝大多数信息检索系统都属于此类型。由于文献信息数量巨大,属性复杂繁多,连贯性强,故文献用单一的独立的记录来描述和表示。记录中含有各种文献特征信息(书目数据)。检索的目的是要获得书目引文和原文出处,而不是现实的数据。也就是说,它通常不必检索某一存入记录中的

各个成分,而只是检索一般的引文。不过,为了鉴别某一特定的引文,有必要标明这些记录的主题内容(即对文献进行主标引)。所以,内容分析工作对文献检索系统来说是最重要的。其次,由于许多供文献检索用的文档往往包含有数十万甚至数百万个记录,所以,文献检索系统中的许多工作都耗费在建立索引结构上面,为的是对检索提问作出快速准确的响应。大型商业性的文献检索系统通常都是专门开发的,检索功能较强,既适用于书目数据和文本信息的存储与检索,也可以用于事实和数值性数据的存取。文献检索系统主要面向文献资源的管理者和利用者,其用户主要是科研和教学人员。此外,它还能对各种管理和决策活动提供有效的信息支持。

### (2) 数据库管理系统

数据库管理系统(Data Base Management System,简称 DBMS)是为任何数据库的建立、操作和维护而设计的计算机软件。它最显著的特征是:存储的数据具有确定的结构,即把数据或信息存储在结构化的数据文件中,系统只能使用那些结构完全确定的数据项。这是它与文献检索系统的主要区别。在 DBMS 中,数据文件通常表示为某种二维表格和不同表之间的关系指示。这种表的每一行表示文件中的一个记录或实体,每一列标明实体的某一属性,根据属性值来区分不同的记录或实体。建库过程中的数据内容分析工作是通过列出有关某实体的一批预定的属性值(如名称、年龄、记录地址等)来迂回达到的。检索时只能查出原先已存入的数据。系统必须能满足用户从不同途径(属性)查询文件中的记录的需要。

DBMS 主要面向数据查询和基层事务管理。其主要用户是行政职员或事务管理员。目前市面上出售的 DBMS 一般具有较好的通用性和可移植性,可供企事业单位直接用来建立事务处理系统(Transaction Processing System)。

### (3) 自动问答系统

自动问答系统(Question-Answering System,简称 QAS)是一种能直接回答人们提出的具体问题的一种计算机系统,曾称事实检索系统,目前已归入专家系统(Expert System,简称 ES)的范畴。问答系统一般由知识库和推理机制组成。

目前,问答系统仍处于实验性使用阶段,只能在某个有限的较狭小的主题领域内采用经过简化的自然语言处理规则来实现。而且,在某个特定领域中采用的语言分析方法通常不能移植到另一些领域。近年来,有人在努力发明某些普遍性的记忆结构和通用的语言处理技术。

问答系统的主要用途是支持或代替专家决策,为知识的采集、存储和利用服务。它采用的知识库技术和自然语言处理技术对文献检索系统和 DBMS 有潜在的应用价值,如应用于自动标引、自动摘录、智能接口设计、自然语言提问处理等方面。可以说,问答系统是信息检索系统的高级形式。在使文献检索系统与问答系统更紧密地结合起来方面也取得了某些进展。

### (4) 管理信息系统

管理信息系统(Management Information Systems,简称 MIS)是建立在数学方法(统计理论、计算方法)、管理科学和信息技术的基础上,对各种用于管理的信息进行采集、加工、分析、传输和保存的一种人机系统。它以各有关(基层)业务部门的信息系统和事务处理系统为基础,把来自各子系统的数据进行加工分析,如分类、排序、检索、筛选、合并、计算、比较、评价、存储及传递等,最后向管理者提供有关计划、管理、作业控制、预测等方面的信息,一般以文

件报表为主。它的主要功能特征是:利用基本的统计技术去分析各种事务信息,然后为最终用户生成各种状态报告;利用标准的运筹学模型去解释非常结构化的管理问题,然后为最终用户生成决策方案。它主要面向中层管理者和资源管理工作人员。

管理信息系统在 60 年代以前主要应用于会计和统计部门,后来扩大到生产管理和市场营销方面。70 年代以后,随着计算机与通讯技术的发展,MIS 逐渐进入一体化阶段,并迅速应用于社会各个领域。它虽然保留有信息存储和检索功能,但更主要的是一种管理决策工具,面向管理人员而不是一般用户。数据的统计分析已成为它的主要功能。

### (5) 决策支持系统

决策支持系统(Decision Support System,简称 DSS)是管理信息系统的高级形式,主要面向中、高层管理和执行决策,尤其是支持半结构或非结构化决策,为决策者提供一种分析问题、构造决策模型和模拟决策过程及其效果的决策环境。所以,它已不是一般的管理工具,而是决策者的“顾问”或“谋士”。

DSS 通常由数据采集系统、数据库管理系统、模型库管理系统、用户接口系统和系统目录构成。它运用数据库、计算机网络和人工智能等技术,把数据采集、存储、检索、检验、分析综合和显示功能结合在一起,灵活性好,适应性强,响应速度快。它输出的是建议、最优方案、最优解,或是不同方案的比较与排队。DSS 的主要功能特征是:①采集、存储、编辑和检索综合性信息或文献;②便利最终用户洞察非结构化问题和作出判断;③利用统计模型、经济模型和会计模型去分析综合信息,并为最终用户生成观点鲜明的状态报告和预测;④利用专门的管理科学模型去评价各种决策方案,并形成推荐性的策略。决策支持系统目前只能建立在较小的主题领域和面向有限的用户。

## 3. 按应用领域划分

信息检索系统或具有查询检索功能的信息系统几乎在各个领域都有应用,且可能以不同形式出现。

(1)科技信息与图书馆管理领域中有书目信息检索系统,全文检索系统、图书馆自动化系统等。

(2)科学研究与工程设计领域中有科研管理系统、实验数据收集与解析系统、分子信息检索系统、CAD 系统等。

(3)行政管理领域有政务信息系统(含居民登记、户籍管理、社会福利、卫生保健、教育、纳税事务等子系统)、秘书业务管理系统、办公自动化系统、交通管理系统等。

(4)经营管理领域有生产管理信息系统、商品流通管理系统、物资管理系统、专利管理系统等。

(5)财政金融方面有银行信息系统、财务管理系统、电子资金转帐系统等。

## 1.4 信息检索方法

信息检索所采用的方法有下列 3 种:

### (1) 追溯法

即利用文献末尾所附的参考文献进行追溯查找的方法,这是一种扩大信息来源最简捷的方法。通过追溯法所获得的文献,有助于对论文的主题背景和立论依据等内容有更深入的理解。但由于是由近及远追溯,年代越远与原文关系越少。为此,美国科学情报所从 1961

年出版了《科学引文索引》(Science Citation Index),中国科学院文献情报中心 1995 年 3 月编出了《中国科学引文索引》(试刊)。

### (2) 常用法

即利用各种检索工具进行查找文献资料的方法。

### (3) 循环法

实际上就是常用法与追溯法的结合,即先利用检索工具查找出一批有用的文献,然后利用这些文献所附的参考文献进行追溯查找。由于参考文献对五年之内的重要文献一般都会引用,根据这个特点,可以跳过利用的五年,然后用检索工具再找出一批文献进行追溯,为此循环交替直至满足检索需要为止。

## 1.5 信息检索的步骤

信息检索可以下面四个步骤进行,即:分析研究课题、选择检索工具、制定检索策略、查阅原始文献。

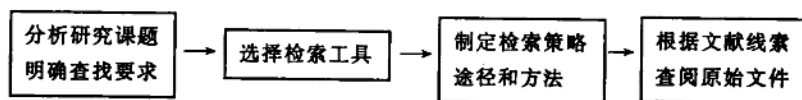


图 1—3 信息检索步骤示意图

### 1.5.1 分析研究课题,明确查找要求

(1)明确研究课题所需的信息内容、性质、水平等情况。比如,是要取得具体的文献资料,还是要掌握某一地区或国家对某一问题发表过的文献资料;是要查找某一年限内对某一问题发表过的文献资料,还是要获得有关某一问题的全部文献资料;等等。

(2)在分析的基础上形成主题概念,包括所需信息的主题概念有几个?概念的专指度是否合适?哪些是主要的?哪些是次要的?等等,力求检索的主题概念准确反映检索需要。

(3)根据检索主题概念的学科性质,确定检索的学科范围。学科范围越具体越有利于检索。

信息检索,实际上是一种检索逻辑的思考与推理,其目的是要掌握检索的必然性,排除检索的偶然性,避免瞎碰瞎撞。

### 1.5.2 选择检索工具书

选择检索工具书主要方法有两种:一是浏览图书馆检索工具室陈列的全部检索书刊,从中挑选确定最为合适的检索工具;二是通过查阅国内外出版的检索工具的介绍,挑选、确定检索工具书。前者的优点是直接具体,缺点是不够全面,因为一般图书馆收藏的文摘刊物毕竟有限。后者的优点是系统全面,提供了查找方法的指导,缺点是缺藏部分无法明确判断其是否适用。

检索的效果常取决于人们对检索工具书的熟悉和了解的程度。因此在查找文献资料之前,必须了解哪些检索工具中收录了与所查专题有关的文献资料,哪些检索工具中该专题的

文献资料比较丰富,哪些检索工具中选录的文献资料质量较高,以及它们报道速度的快慢,分类编排的粗细,摘录质量的高低,附录索引的完备等等。

### 1.5.3 制定检索策略、途径和方法

所谓制定检索策略,就是根据检索分析的结果,制定出一个合理的检索方案,使实行的检索行动能有目的、有计划和有步骤地进行。

(1)根据检索分析,估计哪些图书馆拥有自己所需的信息检索系统。如果本单位、本地区图书馆的信息检索系统不能适应自己的检索需要,那么到外地哪些图书馆、信息所去查找最为合适,以及合理安排先去哪个图书馆,后去哪个图书馆。

(2)在利用检索工具查找所需原始文献前,如有查找图书、专著和述评等一次与三次文献的必要时,就应作出进行检索的合理安排。包括利用何种手段进行检索,检索到以后是否需要经阅读后再进行原始文献的检索,因为一次和三次文献中的参考文献通常提供了许多原始文献的线索。

(3)是否需要通过核心刊物直接检索出几篇所需的原始文献?如果有必要,就要作出怎样进行检索的合理安排,包括了解课题所需要的核心刊物有哪几种,确定检索的次序和大体查找的年代范围等。如果事先已知有关文献的某作者,利用有关检索工具的作者索引,查获该作者所著的文献,然后根据某文献在检索工具中所标引的分类号或主题词,继续查找,即可获得与该文献主题内容相同的其他文献。能否在选好的检索工具中既快又准且全地查获研究课题所需的信息,关键还要熟悉该检索工具使用的检索标识,因为检索的主题词确定不准,即使检索工具中存储了所需文献信息,也是难以查获的。

制定好检索策略后,便以估计最有可能查获的年代范围内先试查1~2年,可能会得到更多的线索,再考虑是否要扩检或缩检。

### 1.5.4 根据文献线索查阅原始文献

通过以上的查阅,如果确认所得的线索有一定的参考价值时,需要进一步了解和详细查阅原始文献资料,可由近而远地查找。

当科技人员花了很多时间仍查不到所需文献时,可以求助于图书馆、信息所的参考咨询人员。在较大的图书馆、信息所,一般都有专职的参考咨询或检索的服务人员,他们的任务就是向读者介绍文献的查找方法,帮助读者解决信息检索中存在的困难和问题。

## 1.6 信息检索语言

要存储信息并使其有序化,就要对其进行标引,检索信息时,也要对检索提问进行标引,从而形成存储标识和检索标识。那么,标引和检索的依据是什么呢?这就用到信息检索语言了。

### 1.6.1 信息检索语言的概念和作用

#### 1. 信息检索语言的概念

信息检索语言是一种专门的人工书面语言,它是信息检索系统存储和检索所使用的共

同语言。其最终目的是为信息检索而用,故名为信息检索语言。

## 2. 信息检索语言的作用

信息检索语言的作用可通过图 1-4 来说明。

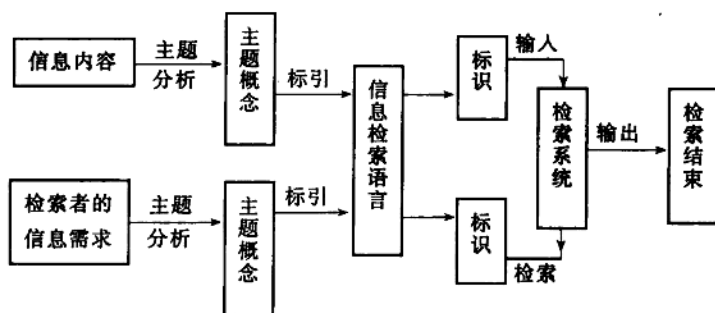


图 1-4 信息检索语言作用示意图

(1)信息检索语言用于标引信息的主题,可把信息的内容及其外表特征简明而有效地揭示出来;

(2)信息检索语言通过对信息的标引,把内容相同及相关的信息集中在一起,加之信息检索语言本身的可序列化,从而可对大量信息进行系统化和组织,使有规律的检索成为可能;

(3)信息检索语言的单一性和表述概念的唯一性以及信息检索语言在信息标引和信息检索时的同时运用,决定了存储标识和检索标识的有效比较,为准确输出所需信息提供了前提。

信息检索语言的作用确立了其在信息检索时的重要地位。由于在反映信息内容与外表特征时的方式和方法的不同,因而形成了信息检索语言的多种类型。

### 1.6.2 信息检索语言的种类

检索语言是用于文献信息的存储外检索的。对于非文献信息,若运用检索语言实际上是把非文献信息转换成为文献信息后再使用。其运用对象虽有差异,但殊途同归。因此,在检索语言中,我们主要介绍文献信息检索语言。

就文献检索语言而言,它反映的是文献信息的内容特征和外表特征。在这些特征中,外表特征如书(篇)名、著者、号码等,表述外表特征的检索语言简单易懂,其形成的又非主要检索途径,故在检索语言中不予专门论述。而表述文献信息内容特征的检索语言,由于其经常使用而又非常重要,故通常所说的检索语言一般均指表述文献信息内容特征的语言。

#### 1. 按组配方式划分

(1)先组式检索语言。指在检索之前,表述文献主题概念的标识已经事先拟定好了的检索语言。先组式语言又分为定组型和散组型两种。定组型是指表达文献主题概念的标识,在编制检索语言词表时已经预先固定组配;散组型是指表达文献主题概念的标识,在编制检索语言词表时并不预先固定组配,而是在标引文献时,根据文献信息的主题内容予以组配,检索阶段也通过相同的组配与之匹配。后面在主题检索语言中讲的标题词语言属于先组定