



高等院校创新精品规划教材

信息检索

主 编 陈明兵

副主编 杜新中 李 理 张胜蓉



中国水利水电出版社
www.waterpub.com.cn

21 世纪高等院校创新精品规划教材

信息检索

主 编 陈明兵

副主编 杜新中 李 理 张胜蓉



中国水利水电出版社
www.waterpub.com.cn

内 容 提 要

本书系统地阐述了信息检索的基本理论与方法、文献信息资源检索；网络资源的检索；中文数据库资源检索与利用；特种文献信息检索；专业信息检索；信息检索综合运用等。同时书中还提供了必要的检索实例和实验。

本书内容新颖、易于掌握、实用性强，理论与实践有机结合，既可作为高等院校本科生、专科生的信息素养教育教材，也可作为从事教学、科研、管理及其他一切与信息检索相关工作的广大读者的自学参考书和工具书，还可作为普通读者查询信息时的学习指导书。

本书配有电子教案，读者可以从中国水利水电出版社网站和万水书苑免费下载，网址为：<http://www.waterpub.com.cn/softdown/>和 <http://www.wsbookshow.com>。

图书在版编目（C I P）数据

信息检索 / 陈明兵主编. — 北京：中国水利水电出版社，2011.8
21世纪高等院校创新精品规划教材
ISBN 978-7-5084-8712-0

I. ①信… II. ①陈… III. ①情报检索—高等学校—教材 IV. ①G252.7

中国版本图书馆CIP数据核字(2011)第115740号

策划编辑：寇文杰

责任编辑：张玉玲

封面设计：李 佳

书 名	21 世纪高等院校创新精品规划教材 信息检索
作 者	主 编 陈明兵 副主编 杜新中 李 理 张胜蓉
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	北京蓝空印刷厂
规 格	184mm×260mm 16 开本 16.25 印张 406 千字
版 次	2011 年 8 月第 1 版 2011 年 8 月第 1 次印刷
印 数	0001—3000 册
定 价	29.00 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

进入新世纪以来,信息检索已经成为提高人们信息素质和综合素质的重要手段之一。随着信息社会化,以及信息技术和信息资源的迅速发展,信息检索的重要性越来越突出。它既是一个平台、一种工具,又是一个“引擎”,通过对“信息检索”课程的教学和学习,使每个学生都善于检索和充分利用信息资源,实现信息和知识效用的最大化。信息素养将成为现代人才的必备条件之一,爱因斯坦曾经说过“用专业知识教育人是不够的”,为此,在大学素质教育中,信息检索已成为当代大学生必须具备的素质,同时培养学生的信息素养已成为科学素养教育的重要内容。“信息检索”课程正是以此为目的而开设的。

本教材是我们从事信息检索教学、科研服务和实际工作的基础上,结合学院专业设置,以及各大检索系统的发展变化,吸收同类教材的优点,组织编写而成的。本教材力图紧扣信息时代的脉搏,突出现代网络信息的检索与利用,注意传统检索与现代检索并重,并针对学生的实际情况,着重介绍学习研究中经常使用到的检索方法,突出实用性。在介绍相关信息检索知识时,采用理论结合实践操作的模式,便于学生举一反三,触类旁通。本教材具有内容新颖、实用性强等特点,在文字上力求简练,在取材上力求新颖、翔实。

本书共 7 章,包括信息检索概述、文献信息资源的检索、网络信息资源的检索、中文数据库资源检索与利用、特种文献信息检索、专业信息检索、信息检索综合运用等内容。

本书由陈明兵任主编,杜新中、李理、张胜蓉任副主编,龚轩涛、喻琼、谢东峰、张丽、薛林、刘兵等参加了部分章节的校审。

在编写本教材过程中,我们参阅了大量文献,其中也包括一些网络资源,因篇幅所限,仅列出主要参考文献,请有关作者谅解,并借此机会向这些文献的作者一并致以诚挚的谢意。

信息检索是一个飞速发展的学科,尽管我们力争全面系统地介绍最新的成果,但由于水平所限,书中内容难免有疏漏和不当之处,恳请同行专家和广大读者批评指正。

编 者

2011 年 4 月

目 录

前言

第1章 信息检索概述	1	2.2.6 手册	35
1.1 文献及其分类	1	2.3 检索工具书	37
1.1.1 基本概念	1	2.3.1 书目	37
1.1.2 文献的分类	2	2.3.2 索引、文摘	42
1.1.3 科技文献	5	2.4 核心期刊	46
1.1.4 图书馆的信息资源构成	8	2.4.1 核心期刊的定义及其特点	46
1.2 信息检索基本原理	9	2.4.2 核心期刊的确定方法	47
1.2.1 信息检索的涵义	9	2.4.3 国内介绍核心期刊的主要工具书	48
1.2.2 信息检索的基本原理	9	2.4.4 我国新闻传播学核心期刊的认定	49
1.3 信息检索语言	11	2.5 文献综述	50
1.3.1 检索语言	11	2.5.1 文献综述简介	50
1.3.2 主要检索语言	11	2.5.2 注释和参考文献	52
1.4 信息检索工具	12	2.6 表谱、图录、名录	55
1.5 信息检索的途径、方法与步骤	14	2.6.1 表谱	55
1.5.1 检索途径	14	2.6.2 图录	56
1.5.2 检索方法	15	2.6.3 名录	56
1.5.3 检索步骤	16	本章小结	57
本章小结	17	习题二	57
习题一	17	第3章 网络信息资源的检索	58
第2章 文献信息资源的检索	18	3.1 搜索引擎	58
2.1 文献信息检索概述	18	3.2 主要的中文搜索引擎	62
2.1.1 文献信息的含义	18	3.2.1 百度 (http://www.baidu.com)	62
2.1.2 文献信息的类型、特征与功用	19	3.2.2 中文雅虎 (http://gbchinese.Yahoo.om)	63
2.1.3 文献信息检索	20	3.2.3 搜狗 (http://www.sogou.com)	65
2.1.4 图书分类	20	3.2.4 新浪网 (http://www.sina.com.cn)	68
2.2 参考工具书	24	3.2.5 网易 (http://www.163.com/)	72
2.2.1 字典、词典	24	3.2.6 天网搜索 (http://maze.tianwang.com/)	74
2.2.2 类书	27	3.3 主要的英文搜索引擎	77
2.2.3 政书	29	3.3.1 Yahoo! (http://www.Yahoo.com)	77
2.2.4 百科全书	30		
2.2.5 年鉴	33		

3.3.2	Altavista (http://www.altavista.com)	80	4.7.7	其他中文数字图书馆	155
3.3.3	Excite (http://www.excite.com)	82	4.7.8	中文科学引文索引	155
3.3.4	Lycos (http://www.Lycos.com)	82	本章小结		155
3.3.5	Google (http://www.google.com)	85	习题四		156
3.4	网络资源检索实例	89	第5章 特种文献信息检索		157
3.4.1	检索实例 1	89	5.1 科技报告的检索		157
3.4.2	检索实例 2	93	5.1.1 国内科技报告的检索工具		157
3.4.3	一般问题举例	97	5.1.2 国外科技报告的检索工具		158
3.4.4	主要搜索引擎的常用高级检索指令	98	5.1.3 因特网上的科技报告检索站点		160
本章小结		100	5.2 会议文献的检索		161
习题三		100	5.2.1 国内会议文献的检索工具		161
第4章 中文数据库资源检索与利用		102	5.2.2 国外会议文献的检索工具		162
4.1 CNKI 数据库系统		102	5.2.3 因特网上的会议文献检索站点		166
4.1.1 初级检索		103	5.3 专利文献的检索		168
4.1.2 高级检索		112	5.3.1 国内专利文献的检索工具		169
4.1.3 专业检索		116	5.3.2 国外专利文献的检索工具		173
4.2 万方数据资源系统		120	5.4 标准文献的检索		184
4.2.1 万方数据资源子系统		120	5.4.1 印刷型标准文献的检索工具		184
4.2.2 快速检索		122	5.4.2 电子型标准文献检索工具		
4.2.3 论文高级检索		127	(数据库系统)		186
4.3 维普资讯网		128	5.5 毕业论文的检索		188
4.3.1 文章搜索		129	5.5.1 国内毕业论文数据库		188
4.3.2 分类导航		134	5.5.2 国外 UMI 的毕业论文数据库		
4.3.3 期刊导航		135	(http://wwwlib.umi.com/dissertations)		190
4.4 超星数字图书		139	5.6 国际联机检索		197
4.5 书生数字图书		144	5.6.1 联机信息检索概况		197
4.6 龙源期刊网		145	5.6.2 DIALOG 系统简介		199
4.6.1 期刊		145	5.6.3 建立 DIALOG 联机		199
4.6.2 电子期刊		146	本章小结		200
4.6.3 龙源期刊网		146	习题五		201
4.7 其他中文数据库		148	第6章 专业信息检索		202
4.7.1 国家科技图书文献中心数据库		148	6.1 文学文献信息检索		202
4.7.2 CALIS 中文资源数据库		149	6.1.1 文学文献信息计算机检索		202
4.7.3 CASHL 资源与服务		149	6.1.2 文学文献信息中文工具书检索		203
4.7.4 免费中文科技文献数据库		150	6.1.3 文学文献信息外文工具书检索		205
4.7.5 国内上网图书馆		151	6.2 艺术文献信息检索		206
4.7.6 特种文献数据库		152	6.2.1 艺术文献信息计算机检索		206

6.2.2 艺术文献信息工具书检索	207
6.3 历史文献信息检索	214
6.3.1 历史文献信息计算机检索	214
6.3.2 历史文献信息工具书检索	214
6.4 百科知识信息检索	219
6.4.1 百科知识信息计算机检索	219
6.4.2 百科知识信息工具书检索	220
本章小结	222
习题六	222

第7章 信息检索综合运用	223
本章小结	236
习题七	236
附录1 利用信息资源进行论文写作和 课题查询	238
附录2 百度搜索技巧	246
附录3 《信息检索》实习报告	252
参考文献	253

第1章 信息检索概述

1.1 文献及其分类

1.1.1 基本概念

(1) 信息。生物及具有自动控制系统的机器通过感觉器官和相应的设备与外界进行交换的一切内容。

在信息社会中,知识和技术是非常重要的—种财富。近年来,信息—词充满了整个社会。广义的信息是物质和能量存在和运动状态的表征。信息并非事物本身,而是事物发出的体现它存在和运动状态的信号和消息。狭义的信息限于人类社会范围内,是人类社会共享的一切知识、学问以及从现象中提炼出来的各种消息的总和,即人们常说的“现代信息”。

(2) 知识。人类对各种大量信息进行思维分析,加工提炼,并加以系统和深化而形成的结果。

知识是人类在改造客观世界的实践中所获得的认识和经验的总和。人们认识客观事物的过程就是人脑对外界事物传来的信息进行加工的过程,而认识飞跃的结果即为知识。就是说,信息被有选择地收入到人的思维系统,经过人脑的贮存、甄别、加工、处理、转换之后便形成知识。知识是信息的一部分,人类不仅要通过信息感知世界、认识世界和改造世界,而且要将获得的信息转变成知识,并当作认识和改造世界的武器。

(3) 情报。是一个发展中的概念。20世纪60年代,情报是特定时间、特定状态下,给特定的人提供的有用知识。70年代,情报是决策所需的知识和智慧。

情报是知识的一部分,是指那些被人们用来解决特定问题所需要并进行交流的知识。情报是由分散的知识经过针对性选择、编码等有序化的加工的产物,它需要经传递才可发生效用。所以,情报具有三性,即知识性、传递性、效用性。由此可以知道,情报不仅取其知识内容,同时强调内容的最佳效应。情报传递或交流除了利用各种类型的文献外,还可通过人的直接对话等方式进行。因此,情报对知识的传递比文献更为广泛。

(4) 文献。以文字、图形、符号、声频、视频等手段记录和传播知识的载体。

习惯上又称“资料”或“文献资料”,包括历史文献和现代文献。文献是人类从事生产实践和科学实验所获得知识的记录,是人类精神文明的结晶。文献作为科学技术的基本要素和信息、资讯、情报交流的最重要载体,对于促进经济建设、发展科学技术、建设社会精神文明与物质文明的重要作用已为越来越多的人所认识。文献记录反映着科学技术、社会发展的成就和水平,积累着人类由生产实践和科学实验中概括和总结出来的大量的科学理论、事实与数据,是取之不尽、用之不竭的最重要的情报源。

文献的三个要素:知识、载体、记录方式。

文献的属性:知识信息性、客观物质性、人工记录性、动态发展性。

科技文献：记录有科学技术文献知识的载体。

信息、知识、情报、文献之间是这样转换的：事物发生发出信息，信息经人脑加工变为知识，知识被记录形成文献，文献经传递应用变为情报，情报应用于实践产生新的信息，知识经传递应用变为情报，情报经过一定时效后，又可还原为知识，情报经提炼、升华还可成为新的知识。四者的关系如图 1-1 所示。

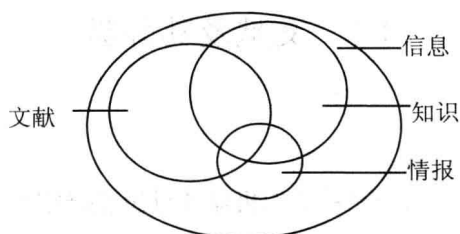


图 1-1 信息、情报、知识、文献的相关性

1.1.2 文献的分类

1. 按出版类型划分（十大情报源）

（1）图书。

“图书”二字，有“图”和“书”两个含义。“图”表示绘画，“书”表示记录文字。我国国家《文献著录总则》和《普通图书著录规则》中，称报纸、杂志等定期出版物以外的非定期出版物为图书，包括单行本、汇编本、丛书本等。联合国教科文组织《关于印刷品统计》文件规定，49 页以下称小册子，这是国际间图书资料的通用统计标准。

图书是一种比较系统完整而又成熟的定型出版物，包括阅读性图书（科学、生产专著、科普与通俗读物、文学艺术著作、教材与教科书）和参考工具书（百科全书、大全、年鉴、手册、辞典、指南）等。

特点：内容系统、全面、成熟、可靠；但编辑出版时间过长，传递情报的速度太慢，包含的内容一般只反映 3~5 年以前的研究水平。

（2）期刊。

期刊又称杂志，有统一的刊名，是采用定期或不定期的方式连续发行的一种出版物。有连续的年、卷、期次编号，其内容多数是由许多短篇篇文章编辑而成的。

期刊按学科专业可分为：综合性和专业性；按内容分：学术性、技术性、行业性、检索性、资料性、科普性、宣传性等多种。

特点：与图书相比数量大、品种多、内容丰富；出版周期短、报道速度快、发行流通面广；能及时反映世界科技发展水平，其中科技期刊是利用率最高的文献类型，占科学家和专家利用全部科技情报的 70% 左右。

（3）科技报告。

科技报告也称技术报告、研究报告。多数是课题研究进展情况的总结记录，反映某项研究的阶段成果或最后的总结报告。其中包括研究方案的选择与比较、成果与失败的记录，常常还附有大量图表及原始数据等。现在世界各国每年都出版大量的科技报告。

特点：科技报告单独成册，页数不等，有机构名称和连续的报告编号，有一定的保密性；能反映新兴科学和尖端科学的研究成果，内容新颖，专业性强，能反映一个国家的研究水平。

如世界著名的美国四大报告，即美国政府的 PB 报告、军事系统的 AD 报告、国家宇航局的 NASA 报告、能源部的 DOE 报告。

(4) 会议文献。

会议文献是指在各种国际学术学议和各国国内重要学术会议上所交流的学术论文、报告及讨论记录等有关文献。会议文献是科技交流的重要渠道，很受科技人员的重视。会议文献分会前或会后出版，正式的会议文献出版时间一般较长，有的长达二三年。

特点：传递情报比较及时、传递情报针对性较强；兼有直接和文献交流两种方式的长处。

包括：会前文献（预印本：是指开会前经选定的论文、报告、材料，由作者单独印成单行本，提交会议参加人员）和会后文献（会议录，是指选择会议的学术报告、交流论文、文献资料及讨论纪要汇编成册后出版发行的资料文献）。

(5) 专利文献。

专利是为了促进技术上发明创造的一种法律制度。它规定在符合专利法的范围内，凡属技术上的发明、创造和革新，经发明人申请并审查属实，由政府公布予以保护，在一定期限内只准发明人专营获利，不准他人抄袭仿造，故称之为专利。

专利文献的主体是专利说明书，是指各国专利局及国际性专利组织在审批专利过程中产生的官方文件及其出版物的总称。

特点：新颖性、创造性和实用性。

包括：专利说明书、申请书、专利分类表、专利索引、专利报道、专利法律文件和专利检索工具等。

(6) 标准文献。

国际标准化组织（ISO）给“标准”下的定义是：标准是经公认的权威当局批准的一个标准化的工作成果。技术标准主要是对工农业产品和工程建设的质量、规格及其检验方法等所作的技术规定，是从事生产、建设的共同技术依据。它是在各有关方面协作下，经过反复实践、选优、协调、统一、规范化等过程而制定下来的，作为一种规章性的技术文献，具有一定的法律约束力。技术标准文献能反映一个国家的经济政策、技术政策、生产水平、加工工艺水平、标准化程度、自然条件、资源分布及利用等情况，对于全面了解一个国家的工农业发展、生产标准化程度，有重要的参考作用。

标准文献是一种规范性的技术文件。它是在生产或科研活动中对产品、工程或其他技术项目的质量品种、检验方法及技术要求所作的统一规定，供人们遵守使用。标准文献包括：国际标准（如 ISO 标准、IEC 标准）、区域性标准（如欧盟标准）、国家标准（如中华人民共和国国家标准）、部标准（如铁道部标准）、企业标准五大类。

(7) 毕业论文。

毕业论文是高等学校或研究机构的学生为获得某种学位，在进行科学研究后撰写的学术论文。一般不在市场上销售，属于非卖品，具有较高的学术水平和参考价值，存放在图书资料部门保存，提供读者查询。一般有学士论文、硕士论文、博士论文。

(8) 政府出版物。

政府出版物是由各国政府部门及其所属机构所发表或出版的文件和资料。

包括：行政性文件和科技文献两类。行政性文件如政府报告、法令、方针政策和政府公布的统计资料等；科技文献如科技报告、技术改革、技术政策、调查报告、科普资料等。科

技文献在政府出版物中占有较大比重，它是了解一个国家的科学技术、各项科技政策及其演变过程等情况的重要文献信息源。

特点：具有正式性和权威性。

(9) 产品资料。

产品资料是指国内外各厂商为推销产品而印发的企业出版物。它记录和反映已投产的产品，故其技术成熟、数据可靠、图文并茂、直观性强，是生产和科研单位研究、分析各国产品技术发展情况和产品水平的重要资料，也是经贸部门判断某项产品价值或洽谈订购产品的重要依据。由于产品不断更新，旧产品不断被淘汰，而代之以新产品样本，因此使用时切不可忽视产品样本的时间性。

包括：产品样本、产品目录、产品说明书、厂商介绍、技术座谈资料等。

(10) 技术档案。

技术档案亦称科技档案，它是记述和反映一个单位或部门在基本建设、生产技术和自然科学研究等活动中具有保存价值并且按照归档制度作为真实历史记录集中保管起来的技术文件资料。它是生产建设与科学技术研究工作中用来积累经验、吸取教训和提高质量的重要文献，具有较高的利用价值。

特点：具有保密性质和内部使用。

包括：各种技术文件、图纸、图片、原始技术记录等资料。

2. 按文献的性质和加工深度与用途划分

(1) 一次文献 (Primary Document)。

一次文献也称原始文献或第一手资料，是以作者本人的研究或研制成果为依据而创作（或撰写）的原始文献。如期刊论文、专著、研究报告、会议文献、档案等。文献检索的最终目的就是要把最适用的具体的一次文献查找出来。

特点：一般较具体、详尽和系统化，内容具有新颖性和创造性。

(2) 二次文献 (Secondary Document)。

二次文献也称检索性文献，是对一次文献进行整理、加工、提炼和压缩之后所形成的文献。如目录、索引、文摘等检索工具。学习文献检索的方法主要是学习二次文献的利用方法。

(3) 三次文献 (Tertiary Document)。

三次文献也称参考性文献，是在一、二次文献的基础上（利用二次文献，选用一次文献），通过广泛深入的分析研究之后，综合概括编写出来的文献。人们常把这类文献称为情报研究的成果，如综述、述评、学科年度总结、文献指南、词典、手册、百科全书、年鉴等。

(4) 零次文献。

零次文献是指通过口头交谈、参观展览、参加报告会等途径获取，即未经记录、未形成文字材料或未经正式发表的原始文献，是人们的“出你之口，入我之耳”的口头交谈，是一些不公开发表的实验记录、备忘录、笔记、手稿、内部报告、技术档案、信件等，没有进入出版、发行和流通这些渠道，收集利用十分困难。

上述的零次文献由于没有进入出版、发行和流通这些渠道，收集利用十分困难，一般不能作为我们利用的文献类型。而一次文献、二次文献和三次文献是一个由分散到集中，由无序到有序，由博而精的对知识信息进行不同层次的加工过程。它们所含信息的质和量是不同的，对于改善人们的知识结构所起到的作用也不同。一般来说，一次文献是基础，是最基本

的信息源，是文献信息检索和利用的主要对象；二次文献是一次文献的集中提炼和有序化，它是文献信息检索的工具；三次文献是把分散的零次文献、一次文献、二次文献按照专题或知识的门类进行综合分析加工而成的成果，是高度浓缩知识信息的一种重要文献情报源，它既是文献信息检索和利用的对象，又可作为检索文献信息的工具。

3. 按不同的物质载体形式划分

(1) 印刷型文献。

以纸张为存储介质；以手写、印刷、打字、复印为记录手段。包括铅印、油印、胶印、木版印刷等。

优点：便于阅读和流传。

缺点：存储密度太低、体积大、份量重、储藏要占很大空间、管理较为困难、难以实现自动输入和自动检索。

(2) 缩微型文献。

一般指以感光材料为载体，以缩微照像等光学记录技术，使印刷型文献按比例缩小的文献资料，包括缩微胶卷、缩微平片和缩微卡片等。

优点：体积小、存储密度高、传递方便、节省储藏空间。

缺点：必须借助缩微阅读机才能阅读。

(3) 声像型文献。

声像型文献又称视听型资料或直感资料，它是运用录音、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、科技电影、幻灯片等。它能直接记录声音与图像，给人以直观的感觉，对于科学家从事科学研究、传播情报均有独特的作用。

优点：存储密度高、直观、真切。

缺点：费用较高。

(4) 机读型文献。

机读型文献全称为计算机可读型文献资料。这是近几年出现并迅速发展的一种新型载体文献。它是指通过编码和程序设计将文字和图像转换成二进制数字代码，把文献转换成数学语言和机器语言，输入计算机，存储记录在磁带或磁盘等磁性载体上。阅读时再由计算机将它输出，转换成文字或图像。它能存储大量情报，并以极快的速度从中取出所需的情报。

包括：磁带、磁盘、磁鼓和光盘等。

优点：存储密度高、存储速度快、原有记录可以删改。

缺点：需要先进的技术设备才能阅读，使用费用高。

1.1.3 科技文献

1. 科技文献的功能

科技文献的功能有：知识积累功能、信息传递功能、继承与借鉴功能、成果评价功能、教育与培养功能。

2. 科技文献对科研工作的作用

科技文献对科研工作的作用如下：

(1) 系统地了解相关领域的研究历史、最新动态及发展方向，分析、归纳、整理出学科前沿的研究热点和难点。

(2) 避免低水平的重复工作, 确保科研质量。

(3) 借鉴他人的成功经验, 寻找解决问题的方法和思路。

3. 信息检索的作用

(1) 信息检索在科学研究中的作用。科学研究作为人类的一种社会活动, 目的是认识自然、社会及人类思维的一般规律, 它是一种创造性劳动, 同时具有连续性和继承性的特点。马克思曾经说过: “这种劳动部分地以今人的协作为条件, 部分地又以对前人劳动的利用为条件”。而无论是“今人的协作”还是利用“前人的劳动”都必须通过信息交流和利用科学文献才能得以实现。

在科学研究中, 任何一个课题从选题、试验直到成果鉴定, 每一步都离不开查阅科技文献, 掌握有关信息。弄清楚哪些工作前人已经做过、用的什么方法、获得什么结果、有什么经验和教训, 以避免重复, 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, 获得较高水平的结果。例如, 1956 年中国科学院生物研究所等单位联合攻关, 首次在世界上人工合成牛胰岛素。他们在定题前就全面、准确地进行了信息检索, 得知英国牛津大学和美国麻省理工学院对该课题已进行了长达 10 年的研究, 并已掌握了牛胰岛素的分子结构。中国的科学家们在获取有关牛胰岛素分子结构的文献资料的基础上, 进一步分析、研究、实验, 最终人工合成了牛胰岛素, 使人类认识生命、揭开生命奥秘迈进了一大步。

(2) 信息检索在经济活动中的作用。当今的信息时代, 信息资源的开发与利用已成为一个国家综合国力的重要标志。正如资本和劳动一直是工业社会的主要变量一样, 信息和知识已成为信息社会的关键变量。信息和知识成为推动科技、经济、文化教育发展的重要杠杆。经济发展依赖于科技进步, “科学技术是第一生产力”。然而科学技术本身并不等于现实的生产力, 还需要转化成现实生产力。这一转化的媒介主要靠科技信息工作, 靠信息资料。充分利用科技信息的功能, 就会促进科学技术和经济的迅猛发展。第二次世界大战后, 日本的迅速崛起即是有力的证明。

例如, 日本仅用 15 年时间就走完了西方先进工业国家半个世纪的历程, 使其跻身于世界经济强国之列。美国哈佛大学教授、美日关系研究所主任埃兹拉·弗格尔说: “如果只举一个原因说明日本的成功, 那就是不间断地、集体地对知识的追求。日本的领导者和国民各阶层中都把学习和追求知识、信息, 看做毕生最重要的事情。”在市场经济体制下, 更显示出信息的获取和利用对经济发展的至关重要的作用。在市场经济体系中, 竞争是通行的法则。任何个人、企业乃至国家, 要想在竞争中站住脚, 都必须充分掌握足够可靠的信息, 并利用它进行科学决策, 才能在竞争中取胜。“以信息求生存, 以信息求发展”已成为我国当今众多有远见卓识的企业家的共识。

(3) 信息检索在生活中的作用。有人可能说: “我一不搞科学, 二不搞管理, 三不经商, 信息检索可能就与我无关了。”在过去可能是这样, 但是由于网络的出现, 改变了我们的生活。现在有些发达国家和地区已实现了衣食住行等诸多问题都可以在信息网络上解决, 试想一下人们如果不掌握信息检索方法将如何生活呢?

总而言之, 生活在信息社会的人们掌握了信息检索技术并能灵活运用, 就能够在激烈的社会竞争中取胜, 就能提高我们的生活质量, 推动社会的进步。

4. 科技文献的发展状况

(1) 数量大、增长快。随着社会经济等的高速发展, 图书馆信息资源日益增长, 数量巨

大。近年来我国每年出版图书 11 万种以上, 期刊约 8000 种, 有统一刊号的报纸 2000 多种, 专利申请 10 万多项, 专利总量已达 100 万件以上。1997 年出版电子出版物 1025 种, 到 2000 年已出版了 33000 多种电子出版物, 只读光盘复制总量为 753 万张。据估计, 现在因特网上大约有 4 亿个网页、200 万个网址。网页上包含 18 亿比特数据信息、20 亿幅图像信息, 而且这些都在随着时间的步伐不断变化、增长。网络连接着包含牛津、哈佛在内的国际一流大学图书馆, 连接着包括美国国会图书馆在内的众多国家图书馆, 连接着包括世界上 Uncover、UMI、SSCI、SCI 等最具有影响的国际联机数据库, 连接着包括 Dialog 公司等 7500 多家信息服务商, 因特网每天发布 18 万条新的消息。

(2) 内容丰富。图书馆是人类信息资源的宝库, 包罗万象。随着电子出版物、网络信息资源的加入, 图书馆信息内容越来越丰富。丰富的信息资源覆盖了人类知识的各个领域, 既有人文社会科学、自然科学、工程信息技术, 也有各种生活服务、娱乐消遣等信息。1996 年, 就有人估计因特网提供的信息总量在 20TB (即 200 万亿字节) 以上, 远远大于全世界每年出版图书期刊的信息总量。

(3) 类型多样。最初图书馆信息资源只有印刷本和手抄本, 20 世纪 70 年代后, 声像资料和缩微制品在我国图书馆的收藏量逐渐增加, 进入 20 世纪 80 年代, 机读文献开始进入图书馆和情报服务机构, 而 20 世纪 90 年代特别是 90 年代后期, 数据库、光盘、网络资源等在图书馆收藏中发展很快。网络环境下的图书馆信息资源不仅包括传统文献信息资源, 如印刷型的图书、期刊、报纸、档案、图谱等; 视听型的录音带、录像带、胶片等; 缩微型的缩微卡片等, 而且还包括电子出版物如磁带、磁盘等, 以及丰富的网络信息资源, 如电子期刊、网络数据库等。可见, 图书馆信息资源类型多样。

(4) 各具特色。网络环境下的图书馆信息资源不仅类型多样, 而且各类型信息资源各具特色。传统印刷型文献具有历史悠久、成熟定型、数量庞大、流通广泛、可长期保存和随时反复使用、符合阅读者习惯等优点, 但同时也有出版周期长、不便更新、内容时效性差、信息量小、检索不便等缺点。电子出版物具有信息存储密度高、信息量大、占用空间小等优点, 但同时也有系统依赖性强、使用不便等特点。网络信息资源具有类型繁多、时效性强、检索较便利、超文本链接、多媒体、多语种等优点, 同时也有信息无序分散、内容准确性较低、可靠性较差等缺点。

(5) 资源共享。在网络环境下, 每个图书馆只是网络整体的一部分或一个节点, 一方面网络环境为图书馆信息资源共享提供了技术条件, 另一方面网络整体协调性对图书馆的信息资源也提出了更高的要求, 使共享性得以强化。当一个图书馆收藏的信息资源转化为一个个数据库并提供网上信息服务时, 其他图书馆完全可以拿来直接使用, 不必再把相同的资源进行加工上网, 共享避免了信息资源的重复建设。

(6) 分布广泛无序。图书馆信息资源不再局限于本馆内部, 而是分布在不同国家和地区。在自愿的基础上, 因特网通过 TCP/IP 协议将不同的网络连接起来, 对网络信息资源本身的组织管理没有统一的标准和规范, 网络信息呈全球化分布结构, 信息资源分别存储在不同国家、不同地区的服务器上, 不同的服务器采取不同的操作系统及数据结构、字符界面、图形界面、超文本形式等, 缺乏集中统一的管理体制, 从整体上网络信息资源尚处于无序状态。

现代科技的发展, 无论是在规模、速度、深度和广度上, 还是在理论和方法上都是过去所无法比拟的。同时科学技术正朝着学科的相互交叉、相互渗透的趋势发展, 反映到上述科技文献的种种特点, 给获取、评价和利用文献信息带来了一定的难度。

5. 课题中查找科技文献的一般规律

文献类型不同, 编纂的目的不同, 在情报研究中各类文献所起的作用也不相同。熟悉文献的类型及其特点, 按课题侧重点去查找所需类型文献的一般规律如下:

- (1) 掌握课题概况: 一般利用图书、专业辞典、百科全书、年鉴、手册等。
 - (2) 了解党和国家对该课题的方针、政策及有关规定: 利用政府出版物、报纸等。
 - (3) 了解该课题有无开发价值: 利用专利资料查新, 查清以前是否有人研究过、研制过。
 - (4) 通过名人录: 了解本领域有哪些著名的学者、他们供职的单位和主要贡献。
 - (5) 通过机构指南: 了解本领域内重要的科学研究机构、企业和厂家, 以及这些机构的一般情况。
 - (6) 通过专科文献指南: 掌握本领域的基本情况及重要检索工具和核心期刊。
 - (7) 利用产品资料、科技档案: 了解竞争对手对该项技术或产品的开发、研制情况。
 - (8) 通过统计部门的统计资料: 掌握课题研究所需要的各种数据。
 - (9) 利用各种检索工具: 查找文献线索。
 - (10) 查阅资料。
- 基础类课题: 侧重于科技期刊、科技报告、毕业论文、会议录和文集等。
 - 开发研究类: 侧重于期刊、专利、样本、标准、产品说明书等。
 - 技术经济类: 期刊、手册、年鉴、报纸、政府出版物。
 - 攻关项目的研究课题: 重点查专利。
 - 科学实验上的攻关课题: 重点查有关科技报告和各种会议录。

1.1.4 图书馆的信息资源构成

图书馆信息资源包括的内容比较丰富, 主要分为文献信息资源和网络信息资源两大类。在两个大类下面又有不同类型的资源。图书馆信息资源已经成为人们获取信息的一个重要途径。图书馆信息资源的详细构成如图 1-2 所示。

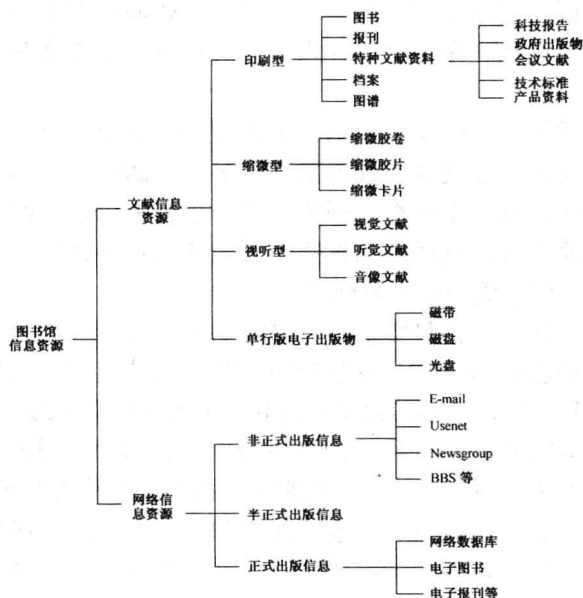


图 1-2 图书馆信息资源的详细构成

1.2 信息检索基本原理

1.2.1 信息检索的涵义

检索的英文是 retrieval 或 search, 其涵义是“查找”, 是指将信息按一定方式组织和存储起来, 并针对用户的需求找出所需要的信息的过程。它又称为“信息存储与检索”(Information Storage and Retrieval)。对于情报用户来说, 信息检索是指查找所需信息的检索过程。一般认为, 信息检索 (Information Retrieval) 包括以下三个方面:

(1) 文献检索 (Document Retrieval)。是以文献为检索对象, 按照一定的检索途径或方法从文献集合中查找出特定的文献, 检索结果是文献或文献线索。它只提供有关的文献, 而不直接回答用户所提技术问题的本身。这类检索数量大、方式不一, 是信息检索的基础部分。

例如“关于自动控制系统有什么参考文献”、“某作家在 60 年代的著述情况”。

文献检索是一种相关性检索, 文献检索是要检索出包含所需要情报的文献。

(2) 数据检索 (Data Retrieval)。是以数据为检索对象, 从已收藏的数据资料中查找出特定数据的过程。如科技数据、金融数据、人口统计数据等。

数据检索是为了满足科技工作者对浓缩信息的特殊需求而出现的, 这种浓缩的信息, 用户可以直接使用, 无需查阅原始文献, 因此可大大节约研究人员的时间, 提高工作效率。

如“查找某种金属的熔点”、“某种材料的电阻”、“喜马拉雅山的高度”等。

数据检索是一种确定性检索, 要直接回答用户所提问题的本身, 提供所需的确切的数据。

(3) 事实检索 (Fact Retrieval)。是以特定的事件或事实为检索对象, 包括事物的性质、定义、原理, 以及发生的地点、时间、前因后果等。检出结果为事实性、知识性的答案。

例如本学年成绩优秀的学生有多少、某同类产品哪种牌号的销量最大。

事实检索也是一种确定性检索, 要求提供用户所需的确定的事实。

数据检索和事实检索的检索结果是“有”或“无”、“正确”或“错误”, 是一种确定性检索。数据检索和事实检索是要检索出包含在文献中的具体情报。

文献检索的检索结果是相关文献或文献线索 (即与某课题有关的若干篇论文、书刊来源出处及收藏地点), 是一种相关性检索。文献检索是要检索出包含所需要情报的文献。

文献检索是三种检索类型中最主要、最基本的形式, 是情报检索中最重要的部分。掌握了文献检索的方法, 就能以最快的速度、在最短的时间内、以最少的精力了解前人和别人取得的经验和成果。

1.2.2 信息检索的基本原理

信息检索是对大量的、分散无序的信息进行搜集、加工、组织、存储, 建立各种各样的检索系统, 并通过一定的方法和手段使存储与检索这两个过程所采用的特征标识达到一致,

以便有效地获得和利用信息源。其中存储是为了检索，而检索又必须先进行存储。图 1-3 给出了信息检索的基本流程。

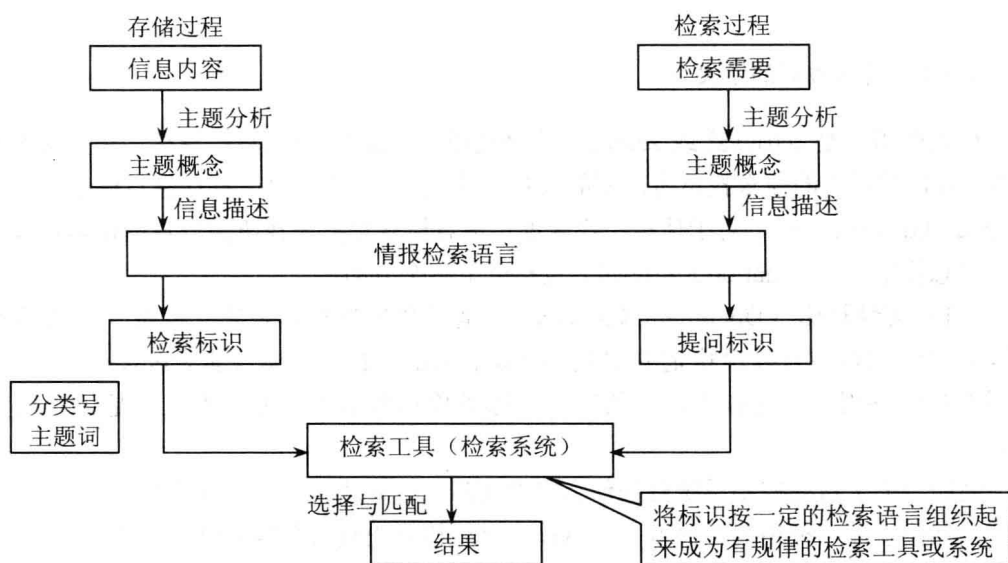


图 1-3 信息检索的基本流程

简而言之，信息怎样存进去，就怎样取出来。这就要求检索的思路必须与存储的思路取得一致。

信息检索的实质是“选择与匹配”的过程。

由信息特征构成的信息检索标志被认为完全反映了信息内容，而由提问特征构成的提问标识则反映了情报用户的需求，因此二者的匹配程度就是信息的相关程度，相关信息就是在某种程度上匹配的信息。图 1-4 给出了信息检索的基本原理。

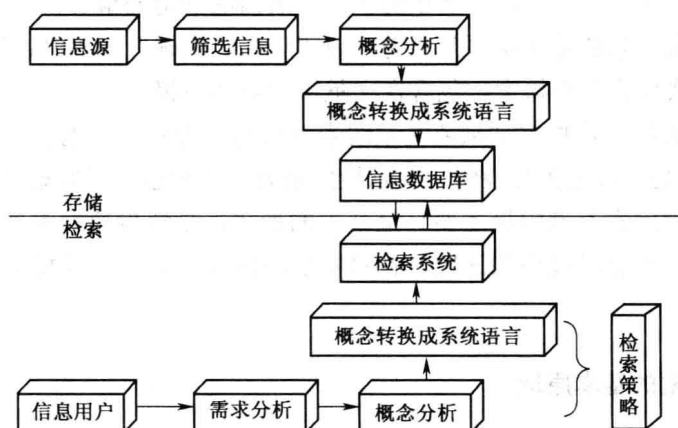


图 1-4 信息检索的基本原理