

信息检索与利用

主 编 赵 兵 吴高魁 李传云

副主编 高文泊 胡友文 伊雪峰



NEUPRESS
东北大学出版社

信息检索与利用

主 编 赵 兵 吴高魁 李传云
副主编 高文泊 胡友文 伊雪峰

东北大学出版社

内 容 提 要

本书针对当前《信息检索与利用》课的教学与实际工作的需要,系统地阐述了信息检索的基本原理和基本方法。介绍了目前国内外一些重要的检索工具和检索方法。全书内容丰富,既介绍了传统的信息检索方法,又用较大篇幅系统介绍了当前国内外流行的检索新技术和新方法。

本书可作为高等院校学生《信息检索与利用》课的教科书,还可作为广大读者信息检索时的参考书。各校使用时可根据本校实际情况予以删减。

图书在版编目(CIP)数据

信息检索与利用/赵兵,吴高魁,李传云主编. —沈阳:东北大学出版社, 2002.8

ISBN 7-81054-778-X

I. 信… II. ①赵… ②吴… ③李… III. 情报检索 IV. G252.7

中国版本图书馆 CIP 数据核字(2002)第 057076 号

出 版 者: 东北大学出版社

(邮编: 110004 地址: 沈阳市和平区文化路 3 号巷 11 号)

出 版 人: 李毓兴

印 刷 者: 沈阳市市政二公司印刷厂

发 行 者: 东北大学出版社

开 本: 787mm×1092mm 1/16

字 数: 427 千字

印 张: 17.125

出版时间: 2002 年 8 月第 1 版

印刷时间: 2002 年 8 月第 1 次印刷

责任编辑: 张德喜

责任出版: 杨华宁

封面设计: 唐敏智

定 价: 25.00 元

垂询电话: 024—83687331 (市场部)

024—83680265 (传 真)

E-mail: neuph@neupress.com

http: //www.neupress.com

前 言

自 1992 年由国家教委《关于印发〈文献检索课基本要求〉的通知》下发以来, 高等学校的“文献检索”课的教学迅速发展并逐渐正规化。随着网上信息的极大丰富, 网络技术的普及与完善, “文献检索”已变为“信息检索”, 检索手段也发生质的变化, 由手工检索、计算机检索发展成光盘检索、网络检索等。

辽宁省教育厅于 2000 年 12 月也下发了《关于印发辽宁省普通高校〈信息检索与利用〉课教学大纲的通知》。为适应新形势的需求, 贯彻、执行这一大纲, 我们组织了省内部分高校的同志编写了本书。其目的在于使学生们能掌握信息检索的基本原理与方法, 了解信息检索在科研立项、研究过程及成果鉴定工作中的作用与意义。培养学生的情报意识与参与习惯, 具有一定的检索技能和查询能力。

本书内容全面、新颖, 注重理论联系实际, 具有通俗、具体、详实的特点。体现出一定的科学性、系统性和实用性。

全书共分八章。内容包括三大部分: 第一部分(第 1—4 章), 检索的基本原理和基本方法。主要介绍信息的基本知识、索引语言和索引方法、科学研究过程与信息检索的关系、检索工具等。第二部分(第 5—7 章), 计算机检索、数据库检索、网络检索。重点介绍国内外一些有代表性且适用面较宽的检索数据库, 搜索引擎等。第三部分(第 8 章), 信息检索思路与情报分析方法。主要介绍信息收集、整理及检索的全过程, 情报分析方法、信息资源开发与利用等。

参加本书编写工作的人员及分工如下:

吴高魁(沈阳工业大学)	1.3 节	3.1 节	3.2 节	3.3 节	4.1 节
伊雪峰(沈阳工业大学)	1.3 节				
金桂花(沈阳工业大学)	4.2 节	4.3 节			
刘 欣(沈阳工业大学)	3.4 节	4.9 节			
刘英煜(沈阳工业大学)	4.6 节				
李传云(沈阳航空工业学院)	第 2 章				
胡友文(沈阳航空工业学院)	1.1 节	1.2 节	3.5 节	4.7 节	4.8 节 4.10 节 4.11 节 4.12 节 4.13 节 4.14 节
刘 涛(沈阳航空工业学院)	1.1 节	1.2 节	3.5 节	4.13 节	4.14 节
陈素清(沈阳航空工业学院)	4.7 节	4.8 节	4.10 节	4.11 节	4.12 节
高文泊(东北大学)	第 5 章, 6 章, 7.4 节				
牟绍斌(东北大学)	4.4 节				
何 洁(东北大学)	4.5 节				
楚晓维(东北大学)	第 8 章				
赵 兵(东北大学)	第 7 章并负责全书统稿				

本教材在编写过程中得到东北大学杨怀教授、中国医科大学郭继军教授、辽宁省科技情报研究所白德良研究员的大量指导和帮助, 在此一并致谢。

由于编者水平所限, 教材中如有缺点和疏漏之处, 恳请读者批评指正。

编 者
2002 年 3 月

目 录

第1章 绪 论	1
1.1 信息的基础知识	1
1.1.1 信 息	1
1.1.2 文献信息源	3
1.2 文献标引及组织方法	5
1.2.1 标 引	6
1.2.2 标引的类型	6
1.2.3 文献标引过程及质量要求	7
1.2.4 分类法与主题法	9
1.3 索引语言与索引类型	10
1.3.1 索引及索引款目	10
1.3.2 索引语言	10
1.3.3 索引的类型	18
思考题	19
第2章 科学研究过程及其与文献检索的关系	20
2.1 科学研究过程	20
2.1.1 科研项目或课题的立项	20
2.1.2 研究、试验、分析阶段	22
2.1.3 科技成果鉴定	24
2.1.4 科技成果奖励	27
2.2 科学研究过程中的信息检索	29
2.2.1 科学研究过程中的信息需求	29
2.2.2 科学研究各阶段的信息检索	32
2.2.3 基础研究和应用研究项目的信息检索	32
2.2.4 发展研究项目的信息检索	34
思考题	36
第3章 中文检索工具	37
3.1 《全国报刊索引》	37
3.2 《中国机械工程文摘》	37
3.2.1 《中国机械工程文摘》概况	37
3.2.2 《中国机械工程文摘》的内容结构	37
3.2.3 《中国机械工程文摘》检索方法	38
3.3 《电子科技文摘》	39

3.4 中国专利文献检索	39
3.4.1 中国专利说明书的种类和编号体系	39
3.4.2 中国专利检索工具	41
3.4.3 中国专利检索方法	41
3.4.4 检索实例	41
3.5 特种文献及其中文检索工具简介	43
3.5.1 科技报告及其中文检索工具简介	43
3.5.2 会议文献及其中文检索工具简介	46
3.5.3 学位论文及其中文检索工具简介	47
3.5.4 标准文献及其中文检索工具简介	49
3.5.5 专利文献及其中文检索工具简介	53
思考题	60
第4章 外文检索工具书	61
4.1 美国《工程索引》(EI)	61
4.1.1 《工程索引》概况	61
4.1.2 《EI》编排结构	61
4.1.3 《EI》检索方法	64
4.2 美国《科学技术会议录索引》(ISTP)	65
4.2.1 《科学技术会议录索引》概况	65
4.2.2 《ISTP》的结构及著录格式	66
4.2.3 《ISTP》检索方法与检索实例	69
4.3 美国《科学引文索引》(SCI)	71
4.3.1 《科学引文索引》概况	71
4.3.2 《SCI》的结构及著录格式	71
4.3.3 《SCI》检索方法与检索实例	75
4.4 《金属文摘》(MA)	76
4.4.1 《MA》概况	76
4.4.2 《MA》月刊文摘本的编排及著录	78
4.4.3 《MA》索引本的编排及著录	81
4.4.4 《Alloys Index》与《Steel Alert》简介	83
4.4.5 检索注意事项与检索实例	84
4.5 《科学文摘》(SA)	86
4.5.1 《科学文摘》概况	86
4.5.2 INSPEC 的产品及服务	87
4.5.3 文摘型出版物	87
4.5.4 《PA》,《EEA》,《CCA》的编排结构及著录格式	88
4.5.5 检索途径与实例	95
4.6 美国《化学文摘》(CA)	97
4.6.1 《化学文摘》概况	97
4.6.2 《化学文摘》的结构与内容	98
4.6.3 《化学文摘》索引的查用方法	100
4.6.4 《化学文摘》的检索方法	104

4.7 美国《国际航宇文摘》(IAA)及其使用方法	105
4.7.1 《国际航宇文摘》概况	105
4.7.2 《IAA》的内容编排及著录格式	106
4.7.3 《NASA 叙词表》	109
4.7.4 《IAA》的使用方法与检索实例	110
4.8 科技报告的检索工具及其使用方法	111
4.8.1 概 况	111
4.8.2 《GRA&I》的内容编排及著录格式	111
4.8.3 使用方法与检索实例	113
4.9 德温特专利检索工具	114
4.9.1 德温特出版物概况	114
4.9.2 WPIG 的各种索引格式	116
4.9.3 《文摘周报》的著录格式	119
4.9.4 德温特专利的检索方法	120
4.10 日本《科学技术文献速报》及其使用方法	121
4.10.1 《速报》概况	121
4.10.2 《速报》的内容编排和著录格式	121
4.10.3 《速报》的使用方法和检索实例	123
4.11 俄罗斯《文摘杂志》(PЖ)及其使用方法	124
4.11.1 《文摘杂志》概况	124
4.11.2 《文摘杂志》的使用方法与检索实例	126
4.12 学位论文的检索工具及其使用方法	127
4.12.1 《国际学位论文摘要》(DAI)概况	128
4.12.2 《DAI》内容编排与著录格式	128
4.12.3 《DAI》使用方法与检索实例	129
4.13 标准文献检索工具及其使用方法	130
4.13.1 ISO 标准	130
4.13.2 ISO 标准文献的检索工具及其使用方法	130
4.13.3 《ISO Catalogue》使用方法与检索实例	134
思考题	135
第 5 章 计算机检索概述	136
5.1 计算机检索的发展历史	136
5.2 计算机检索的原理	137
5.3 文献数据库简介	137
5.3.1 数据库的定义和类型	137
5.3.2 数据库的构成	138
5.4 计算机检索中常用的运算符	140
5.4.1 逻辑算符	140
5.4.2 优先算符	140
5.4.3 截词符和屏蔽符	141
5.4.4 位置算符	141

5.5 计算机检索策略的制定及调整	142
5.5.1 概念的选取	142
5.5.2 数据库的选择	143
5.5.3 检索策略的制定及调整	143
思考题	144
第6章 数据库检索系统	145
6.1 Dialog 联机检索系统	145
6.1.1 系统简介	145
6.1.2 Dialog 检索系统的主页	145
6.1.3 Dialog Select 检索界面	145
6.1.4 DialogWeb 检索界面	147
6.1.5 DialogClassic 检索界面	151
6.1.6 DialogLink 检索界面	152
6.1.7 Dialog 检索算符	153
6.1.8 基本检索指令	153
6.1.9 其他检索指令	154
6.1.10 禁用词和禁用字符	155
6.1.11 结果显示格式	155
6.1.12 基本索引和辅助索引	155
6.1.13 Database Bluesheets——数据库蓝页	156
6.1.14 使用 Dialog Database Catalog(分类目录)	157
6.1.15 利用总索引文档选择数据库	158
6.1.16 One Search——一次检索功能	158
6.1.17 检索实例	158
6.2 OCLC FirstSearch 检索系统	162
6.2.1 OCLC 检索系统简介	162
6.2.2 OCLC FirstSearch 的检索界面及检索系统	164
6.3 Ei Compendex Web	169
6.3.1 Ei Compendex Web 简介	169
6.3.2 Ei Compendex Web 的检索界面及使用方法	169
6.3.3 检索结果的显示	175
6.3.4 检索结果的处理	175
6.4 CSA——剑桥科学文摘	176
6.4.1 CSA 概况	176
6.4.2 检索界面及检索步骤	176
6.5 PQDD——博士、硕士论文数据库	182
6.5.1 PQDD 简介	182
6.5.2 检索界面及使用方法	182
6.5.3 检索式的构造	184
6.5.4 检索结果的显示	185
6.5.5 作标记的检索结果的处理	185
6.5.6 重新检索/二次检索	186

6.6 ingenta 检索系统	186
6.6.1 ingenta 检索系统简介	186
6.6.2 检索界面及使用方法	187
6.6.3 检索式的构造	189
6.6.4 检索结果的显示及处理	189
6.6.5 期刊浏览界面及使用方法	190
6.6.6 主题目录的使用	192
6.6.7 查看检索历史、调用检索策略	192
6.6.8 按钮的功能	192
6.7 WOSP-ISTP	193
6.7.1 WOSP-ISTP 简介	193
6.7.2 WOSP-ISTP 的检索界面及使用方法	194
6.7.3 检索式的构造	195
6.7.4 检索结果的显示	196
6.7.5 检索结果的处理	197
6.8 中国资讯行	199
6.8.1 中国资讯行简介	199
6.8.2 中国资讯行——中国教育网镜像站点的使用	199
6.9 万方数据资源系统	201
6.9.1 万方数据资源系统简介	201
6.9.2 科技信息子系统的使用方法	202
6.9.3 数字化期刊子系统的使用方法	205
6.10 中国期刊网全文数据库	207
6.10.1 中国期刊网全文数据库简介	207
6.10.2 登录站点	209
6.10.3 检索界面及使用方法	209
6.10.4 下载全文浏览器	212
6.10.5 全文的处理	213
6.11 Kluwer Online Journals——电子期刊全文数据库	214
6.11.1 Kluwer Online Journals 简介	214
6.11.2 登录系统	214
6.11.3 期刊的检索与浏览	214
6.11.4 检索方式	215
6.11.5 检索式的构造	217
6.11.6 检索结果的显示及处理	218
6.12 ELSEVIER SCIENCE 电子期刊全文库	219
6.13 ProQuest Academic Research Library——全文期刊与报纸数据库	223
6.13.1 简介	223
6.13.2 ARL 数据库的使用方法	223
思考题	228
第7章 INTERNET 信息资源的检索	230
7.1 INTERNET 简介	230

7.2 网上信息检索	231
7.2.1 网上信息的分类及特点	231
7.2.2 Internet 检索信息的基本规则	231
7.2.3 网上信息检索	232
7.2.4 网给信息索引文档	232
7.3 搜索引擎	233
7.3.1 搜索引擎的分类	233
7.3.2 搜索引擎的组成部分	234
7.3.3 搜索引擎的评价	234
7.3.4 搜索引擎的局限性	234
7.3.5 中外著名搜索引擎简介	235
7.4 其 他	241
7.4.1 数字化图书馆	241
7.4.2 国家科技图书文献中心简介	241
7.4.3 专利信息的网上检索	243
7.4.4 科技报告的网上检索	244
7.4.5 标准信息的网上检索	245
思考题	246
第 8 章 信息检索思路与情报分析方法	247
8.1 文献信息的收集、整理和分析	247
8.1.1 文献信息的收集方法	247
8.1.2 文献信息的整理方法	248
8.1.3 文献信息的分析方法	249
8.2 文献信息分析成果的表达	250
8.3 信息资源的开发与利用和综述实例	251
8.3.1 文献信息调研的准备	252
8.3.2 检索策略和部分检索结果	252
8.3.3 文献信息研究综述	253
思考题	254
附录 1 数字化图书馆	255
附录 2 各国国家图书馆	255
附录 3 国内外学会协会网址	256
附录 4 专利网站	257
附录 5 科技报告网站	258
附录 6 标准网站	259
附录 7 在线学术期刊	260
附录 8 中文核心期刊目录	262
参考文献	263

第1章 绪 论

1.1 信息的基础知识

在日常的工作、学习和生活中，人们时刻都要与信息打交道。随着信息社会的到来，信息资源越来越受到人们的重视，人们对信息的认识也逐步加深。而现代科学技术的进步，使人们认识到信息的价值，把信息、物质和能源并列，称之为维系社会发展的要素。

1.1.1 信 息

(1) 信息的定义

信息论创始人、美国科学家申农认为：“信息是在通讯的任何可逆重新编码或翻译中那些保持不变的东西。”

控制论创始人、美国统计学家维纳认为：“信息是人们在适应外部世界，并使这种适应反作用于外部世界过程中，同外部世界进行相互交换的内容的名称”，他还认为，“接受信息和使用信息的过程，就是我们适应外部世界环境的偶然性变化的过程，也是我们在这个环境中有效地生活的过程。”

英国学者阿希贝则提出：“信息是被传输的变异度”。他认为信息的本性在于事物本身具有变异度。

我国学者对信息的概念也有很多研究和阐述。

我国《辞海》载：“信息是客观存在的消息、情况、情报等。”

《新华辞典》载：“信息是指音信、消息；信息论中指用符号传送的报道，报道的内容是接收符号者预先不知道的事物的运动状态和关于事物运动状态的陈述。”

尤以我国学者钟义信对信息及其相关概念的论述最为精辟。

钟义信站在信息科学的高度将信息界定为：“事物运动的状态与方式”。这是广义的信息定义，是最普遍的层次。

从以上论述中可以看出，虽然信息作为新的研究对象内容丰富、涉及领域广泛，但信息始终有三个基本点：信源、信宿、载体。信源指发出信息的物体；信宿指接受信息的物体；载体是承载信息的物体，客观事物的实质内容必须通过一定的载体传递，才能为其他事物所了解、认识，从而成为信息。

(2) 信息、知识、情报、文献之间的区别

信息与知识、情报、文献之间虽有一定的内在联系，但也有所区别。

知识是人类社会实践经验的总结，是人的主观世界对于客观世界的概括和如实反映。也可以说知识是人类通过信息对自然界、人类社会及思维方式与运动规律的认识和掌握，是人的大脑通过思维重新组合的系统化信息的组合。因此，人类要通过信息来感知世界、认识世界和改造世界，又要根据所获得的信息组织知识。知识是信息的一部分，是一种特定的人类信息。人类生活环境中普遍存在的信息是知识的原料，这些原料经过人类接受、选择、处

理,才能成为新的系统化的知识。

情报是指有目的、有时效,经过传递获取的涉及到一定利害关系的特定的情况报道或资料整理的结果,它是一种特定的知识性信息。信息与情报是有区别的。信息的范围比情报广泛得多;情报的传递具有机密性,传递手段有一定的特殊要求;情报是知识的一部分,情报的知识性较信息的知识性强;情报的得失往往伴随着一定的利害关系,而信息的得失则不一定表现出明显的利害关系。

文献在《中华人民共和国国家标准·文献著录总则》(GB3792.1—83)中定义为:“文献,记录有知识的一定载体”。知识、载体、记录是构成文献的三大要素。载体是文献的外部形式,是知识的包装或运载方式,是通过记录形成的物质实体。记录是信息的一种人工编码,通过印刷、光、电和磁等技术手段生成含有知识内容的标识符号,形成多种媒体的文献类型。科技发展的连续性和继承性缘于文献载体记录的科技知识。

(3) 信息的特征

信息不是物质本身,但它来源于物质,信息的物质性决定了它的一般属性。

① 普遍性。世界上任何运动着的事物无时无刻不在生成信息,只要有事物存在,只要有事物在运动,就存在着信息。信息无所不在,物质的普遍性以及物质运动的规律性决定了信息的普遍性。

② 客观性。信息不是虚无缥缈的东西,是现实中各种事物运动的状态与方式的客观反映。由于事物及其状态、特征和变化是不以人的意志为转移的客观存在,所以反映这种客观存在的信息也具有客观性。信息不仅其实质内容具有客观性,而且一经形成且与载体结合,其本身也具有客观实在性。

③ 抽象性。信息本身是看不见摸不着的,我们能够看得见摸得着的只是信息载体(包括语言、文字、图画、符号、纸张、光盘、人的大脑等)而非信息内容。对于认识主体而言,获取信息和利用信息具备抽象能力,正是这种能力决定着人的智力和创造力。信息的抽象性增加了信息认识和利用的难度,并从而对人类提出了更高的要求。

④ 依附性。是抽象性的延伸。信息的记录存贮以及交流和共享必须依附于或借助于一定的物质载体,以某种载体形式表现出来,没有载体就没有信息。

⑤ 可加工性。信息数量庞大,质量高低不一,而人们对信息的需求往往具有一定的选择性。信息价值的发挥也需要进行不同层次的加工处理,由原始信息可以加工成二次信息,二次信息经过分析、研究综合,又可加工成三次信息。每次加工都可改变原有信息的结构。赋予信息新的价值。

⑥ 传递性。信息在运动中产生,在传递中发挥价值。信息可以在时间和空间上进行纵向和横向传递。信息的获取利用以及反馈必须借助于信息的传递。信息传递是通过信道(信息传输通道)进行的。信息系统就是由信源、信道、信宿组成的有机整体。

⑦ 共享性。信息能够通过时空进行传递,因此能够被人类所共享,信息价值的实现需要通过信息的使用。与实物使用不同,同一信息可以同时被两个以上的多个用户使用,而且并不因信息的多人多次重复使用而丢失其内容。正如萧伯纳所举的“苹果与思想”的例子,苹果交换之后交换双方仍然各有且仅有一个苹果,但思想交换之后交换双方都拥有了两种思想。

信息除了上述特性之外还有时效性、动态性、可转化性、可伪性等特性。

(4) 信息的类型

信息分类一般可按其内容和渠道等标准划分。

① 按内容信息可分为主观信息和客观信息。客观信息是不加主观判断的如实客观的信息,它一般全面客观地描述一个问题的各个方面,使人们对问题有一个全面的概念。客观信息有基本的事实和客观的描述,用户可以直接利用这样的信息。

主观信息一般是依据事实和分析说明个人的观点和见解。在对一个事件、论题进行评估时能提供很多有价值的信息。

② 按传播渠道划分信息可分为口传信息、实物信息、文献信息。

口传信息又称准信息,它借助声波和空气口耳相传。口传信息是一种瞬时信息,是非正式信息交流中一种最常见的形式,其特点是速度快,传播范围广泛,但缺乏查考依据。

实物信息泛指作为产品信息进行收集和展示的、经过人类有目的加工的产品,如机器、仪表、服装、书籍等。实物信息直接、直观、真实、可靠。

文献信息是指用文字、图形、符号、声频、视频等手段记录在一定物质载体上的信息。这是一种在空间上和时间内传递信息最有效的手段,是最基本的信息传递方式。本教材所讨论的信息以文献信息为主要对象。

1.1.2 文献信息源

(1) 信息源概述

信息源即信息的来源。信息源可以积累信息,因此信息源和吸收源之间就形成信息位差,也称信息势,信息势的存在是信息流和信息交流活动产生的前提。需要指出的是信息源不等于信息资源,信息源是蕴含信息的一切事物,而信息资源是可利用的信息的集合,是一种高质量、高纯度的信息源。信息源概念与信息概念是紧密联系在一起的,信息有不同的层次和类别,信息源也有不同的层次和类别。

依据信息源的层次及其加工和集约程度,信息源可分为四次信息源:一次信息源也称本体论信息源,所有物质均为一次信息源;二次信息源也称感知信息源,人的大脑所储藏的潜在信息资源是最主要的二次信息源,传播、信息咨询、决策等领域所研究的也是最主要的二次信息源;三次信息源也称再生信息源,主要包括口头信息源、体语信息源、文献信息源和实物信息源四大类型;四次信息源也称集约信息源,是文献信息源或实物信息源的集约化,如图书馆、数据库、各类博物馆、标本室等。其中文献信息源包括印刷型文献信息源和电子文献信息源,是人们获取信息的主要来源。

(2) 文献信息源的类型

文献信息源由于种类繁多,数量庞大,如果不对其进行分类,很难被人们掌握利用。

① 按加工处理的深度划分产生零次文献、一次文献、二次文献、三次文献。

a. 零次文献。是指未形成一次文献之前的信息载体,即未正式公开发表的最原始文献。如未正式发表的书信、手稿、讨论稿、谈话记录等。零次文献虽然没有进入正式的信息交流渠道,但却是新思想、新理论、新产品产生的最初的、最重要的文献记录形式,具有重要的情报价值。

b. 一次文献。通常指作者根据本人的科研和工作成果为基础创作和撰写的文献。它们或多或少包含着作者的创见。一次文献是未经过他人的加工、提炼和压缩,是记录人类知识和智慧的主体,是信息传播的实质部分。

c. 二次文献。是指对一次文献进行整理、加工和系统编排的各种检索型出版物。包括

索引、文摘、目录、题录等检索工具,是人们及时、迅速查找文献线索的工具。

d. 三次文献。是指在利用一次文献和二次文献的基础上,根据用户需要,对特定主题的信息进行收集、选择、分析、综合、浓缩而形成的文献。如综述、述评、分析报告、手册等高智能的加工形成的文献。

从情报学的角度看,一次文献是检索对象和目标,二次文献是检索工具和手段,三次文献是情报研究成果,既可是目标,又可是手段。

② 按载体形式划分,文献信息源可分为印刷型、缩微型、电子型、声像型。

a. 印刷型。是以纸张为介质,以印刷、复印等记录手段产生的媒体形式,是人类使用时间最长、最普及的一种载体形式。由于印刷型文献便于携带、价格低廉、阅读不需辅助设备,因此它是至今为止仍占主导地位的文献类型。缺点是存贮密度低,不便于加工、整理和收藏。并且由于造纸原料越来越少,其价格也会不断上涨。

b. 缩微型。是以感光材料为存贮介质,以缩微照相为记录手段而产生出来的文献形式。包括缩微胶卷、缩微胶片、缩微卡片、缩微印刷品等。其特点是体积小,存贮量大,价格适中,但必须通过缩微阅读机阅读,清晰度不高。

c. 电子型。其前身是机读型,是以磁性材料为介质,以打字、或光学符号等识别装置等高新技术为记录手段形成多种类型的电子出版物。包括磁带、磁盘等。特点是存取速度快、存贮量大、易于改变或更新。缺点是需要相应的技术设备才能阅读、费用较高。

d. 声像型。也称视听型,指通过特定设备,使用声、光、磁、电等技术将信息表现为声音、图像、动画等形式,给人以直观形象的感受。包括唱片、录音带、幻灯片、录像带等。特点是存贮密度高、内容直观真切、表现力强,传播效果好。但也必须借助相应设备才能阅读。

以上几种文献类型中,印刷型是基础,电子型是发展方向,它们将在相当长的一段时间内共存,相互补充,各自发挥优势。

③ 按出版类型划分,文献信息源又可分为十大类。

a. 图书。图书大多是对已发表的科研成果、生产技术和经验或某一知识领域系统地论述和概括。它往往以期刊论文、会议论文、研究报告及其他第一手资料为基本素材,经过作者的分析、归纳、重新组织而编写成的,有的也包含一些未发表过的研究成果和资料。正式出版的图书一般都有 ISBN 号(International Standard Book Number),其定长为 10 个数字,ISBN 号采用弹性存贮的方法,可对全世界的图书实现上千年的管理。

图书一般包括阅览类图书和参考类图书。阅览类图书有教科书、专著等,参考类图书包括辞典、百科全书、手册和年鉴等。图书的缺点是出版周期相对较长,信息传递较慢。如果对范围较广的问题获得一般性了解,或对陌生问题作初步学习,阅读图书比较好,因其知识系统,生命期长。

b. 期刊。是指凡是有一个固定名称、统一的出版形式和装帧,有一定的出版规律,每年至少出一期,每期载有两篇以上不同作者写的论文,按一定的编号连续出版下去的一种出版物。期刊的特点是数量大、种类多、内容丰富、出版周期短(常有周刊、月刊、双月刊、季刊、年刊等)、报道速度快、发行流通面广、连续性强,尤其某专业的核心期刊是研究人员获取信息最常用的来源,具有较高的研究参考价值。期刊也有国际标准编号 ISSN(International Series Standard Number),加以管理。此外期刊还可分市、省(地区)、全国、国际等级别,以通报、学报为名的期刊大多由大学和学术团体出版,学术水平较高。

c. 科技报告。是关于某项科研成果的正式报告或其进展阶段的实际记录。科技报告的出版形式比较特殊,每份报告自成一册,篇幅长短不等,有连续编号,装订简单,出版发行不规则。科技报告内容新颖、详尽、专深,常附有大量的数据、图表、原始实验记录等资料。目前美、俄、英、法、德、日等国每年都发表大量的科技报告,如美国政府著名的四大科技报告:AD报告、PB报告、NASA报告、DOE报告。

d. 会议文献。是科技工作者在学术会议上宣读的论文或报告,包括会前出版物和会后出版物等。会前出版物多难以获取,如会议日程表、会议论文预印本、论文摘要等;会后出版物指会议结束后,经主办单位整理、编辑出版的正式文献,其内容系统、完整。一般有会议录、论文集等形式。会议文献特点是传递信息及时、针对性强、内容新颖、学术水平高和出版形式多样等,是一种十分重要的科技情报源。

e. 专利文献。是指获得国家专利权的、记载发明创造内容的技术文献,包括专利说明书、专利公报、专利检索工具,以及与专利有关的法律文件等。其中专利说明书是专利文献的主体和核心,内容包括发明的详细说明、专利权范围、插图等。其文体通常既含糊又严谨。即为了保护技术关键,在说明技术问题时常含糊其词,而在权力要求部分则采用严格的法律叙述。

f. 学位论文。国际标准化组织给学位论文下的定义是:“著者为取得专业资格的学位而提出的,介绍他的发现与某种结论的调查和研究论文。”一般有学士论文、硕士论文、博士论文。其中硕士论文和博士论文有较高的参考价值。学术论文一般不出版发行,但其理论性较强,并附有大量的参考文献,可看出有关专题的发表过程 and 方向,也是较为重要的科技情报源。

g. 政府出版物。是各国政府部门及其所属机构发表的文件,可分为行政性文件和科技性文件两大类。行政性文件包括国会记录、司法资料、方针政策、规章制度和调查统计资料等。科技性文件包括政府所属部门的科技报告、科普资料和技术决策文件等。政府出版物的出版形式有图书、期刊等,是了解有关国家的科技经济政策和社会经济发展情况的重要依据,有较高的参考价值。

h. 技术标准。是经过公认的权威当局批准的标准化工作成果,采用文件或规定基本单位两种形式固定下来,形成文献,反映当时的技术工艺水平和技术政策。

i. 产品样本。是产品制造商介绍其产品的文献。如产品说明书、产品目录等。其特点是技术成熟可靠、装潢美观、图文并茂,对于技术人员的产品选型和外观设计有很大的参考价值。

j. 科技档案。是科研或生产单位对某工程项目所做的全部有关技术文件进行归档保存的文献。科技档案包括工程图纸、图表、图片、原始记录、任务书、技术合同、审批文件、计划、方案或实施措施等。它是科技储备最高、最完善的文献形式。

1.2 文献标引及组织方法

信息检索系统包括信息的存贮与检索两个方面。信息存贮是指编制检索工具和建立检索系统;信息检索是指利用这些检索工具和检索系统来查找所需信息。而在检索存贮与检索中检索语言起着语言保证的作用,是沟通标引人员和检索人员双方思想的桥梁(如图1-1所示)。

信息存贮时,文献标引人员首先要对各种文献进行主题分析,即把它所包含的情报内容

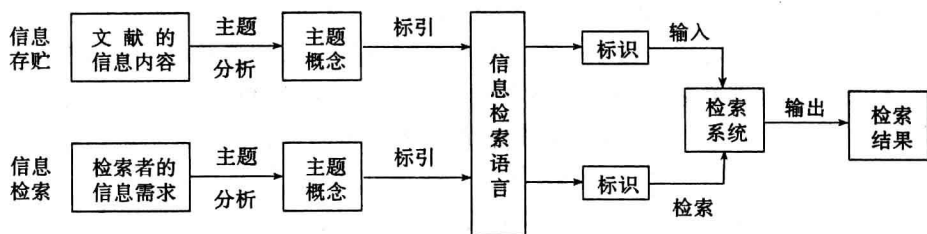


图 1-1 信息存贮与检索系统

分析出来，使之形成若干能代表文献主题的概念，并用检索语言的语词（标识）把这些概念标示出来，然后纳入检索工具或检索系统，此时检索语言即称标引语言；检索信息时，检索人员也首先要对检索课题进行主题分析，即把它所涉及的检索范围明确起来，使之形成若干能代表信息需要的概念，并把这些概念转换成检索语言的语词（标识），然后从检索工具或检索系统中查找正是用该语词标引的文献，从而找到包含所需信息的文献，此时检索语言即称索引语言。索引语言及索引方法将在 1.3 节中介绍，本节主要论述文献标引及组织方法。

1.2.1 标 引

国家标准 GB3860—83《文献主题标引》中对标引的定义是“标引是对文献进行主题分析，从自然语言转换成规范化的检索语言的过程。对文献给予分类号标识的过程，称为分类标引，给予主题词标识的过程，称为主题标引”。所以标引的实质是对文献给予主题词或分类号的过程。文献标引既涉及到文献的外部特征，又涉及到文献的内部特征。

文献只有获得检索标识后，才能按检索标识的逻辑次序加以组织，转化为有序的集合，使从文献内容特征进行检索成为可能。文献标引的目的就是通过标引人员把文献和文献用户联系起来，使用户能在大量的文献中全面、准确、迅速地查到特定的文献。由此可见，没有标引就没有检索，文献标引是建立文献检索系统和进行文献检索的关键环节，文献标引工作也是整个图书情报工作的重要基础。

1.2.2 标引的类型

按不同角度划分可将标引分为不同的类型。

① 按文献检索方法划分，可分为主题标引和分类标引。

标引是描述文献内容的一种形式，是选择最恰当的词汇表达文献内容的过程。对文献给予分类号标识称分类标引；给予主题词标识称主题标引。

② 按操作者划分，可分为人工标引和自动标引。

人工标引又称手工标引，是直接由标引人员赋予文献以某种检索标识的标引。

自动标引又称机器标引，是由电子计算机等机器设备来赋予文献某种检索标识的标引。自动标引又分为自动抽词标引和自动赋词标引。自动抽词标引是由计算机直接从文献（题目、摘要）中自动抽取表达该文献主题的关键词作为检索标识的一种标引；自动赋词标引则是依据文献本身的词汇特点，由计算机将主题词表中的主题词自动赋予文献的一种标引。

③ 按使用检索标识的规范程度划分，可分为受控标引、自由标引、混合标引。

受控标引是用规范化的检索语言，如标题词语言、叙词语言，等等，作为描述和表达文献主题检索标识的标引。

自由标引又称非控标引,是直接使用自然语言中未经规范的自由词作为表达文献主题检索标识的标引。

混合标引又称半控标引,是指标引文献时既使用规范化的主题词,又同时使用自然语言中未规范的自由词、关键词等来描述和表达主题的标引。

④ 按文献标引的适当程度划分,可分为过度标引、过粗标引、适度标引。

过度标引是指标引时,脱离文献的主题需要,标引了过多、过深的主题词,也就是说,给文献标引了一些多余的不反映文献主题的主题词。其后果一是造成人力、物力和时间的浪费;二是影响检索系统的精度,产生大量的检索噪声;三是造成误检,降低了文献的检准率。

过粗标引是指标引的主题概念不能全面反映文献的实际内容,标引过于泛指,其弊病是降低文献检索的检全率。

⑤ 按文献标引的深度和方式划分,可分为整体标引、全面标引、补充标引、重点标引。

整体标引又叫浅标引,即对一部书或一篇文章的整体主题用一个标识来进行概括性标引。

全面标引又叫深标引,即对一篇文献的各个局部主题或构成完整主题的各个主题因素分别标引,必要时可同时对整体主题作概括性标引;如无必要,就不再对整体主题作概括性标引。

补充标引是一种辅助标引方式,即除了对一篇文献的整体主题作概括性标引外,再选择个别重要的局部主题或主题因素作单独标引。

重点标引也称对口标引,即只选择一篇文献中对本单位服务对象有情报价值的个别局部主题作标引。

1.2.3 文献标引过程及质量要求

文献标引是整个文献工作的重要环节,是文献检索的基础和前期工作,标引质量如何将直接影响文献检索系统的质量和性能,进而影响情报工作的开展。因此标引工作作为一项复杂的技术性工作应遵循一定的工作程序。

(1) 查重

查重就是利用工作题目录或机读文档,查明待标的文献与已入藏文献的关系:如属于复本,则仍可使用原文献的分类标识和主题标识;如原来的标引有明显错误,则应予以纠正;如属于新入藏的文献,则按标引程序进行操作。查重不但可减少标引误差,保证文献标引的一致性,也可提高标引速度。

(2) 分析主题,选定主题概念

主题是指文献所研究的对象或问题,即文献的中心内容。文献的主题分析是标引工作的重要环节。主题分析是指对于需要标引的文献首先对文献的内容特征进行分析,在充分了解文献内容及其学科属性、研究对象的基础上,深入分析主题的类型、主题结构及构成要素,对有检索意义的主题概念进行概括、提炼和选择的过程。主题分析时对文献内容的分析不应受检索语言的限制,标引人员可用自然语言对文献主题内容进行描述,且尽量做到全面、充分,不但对各类型主题,包括单主题和多主题、显性主题与隐含主题、主要主题与次要主题等进行分析,同时进行主题结构分析,抓住“对象”、“方面”等主要成分与次要成分,具有独立检索意义和辅助检索意义的主题要素,以提高主题分析乃至文献标引的质量。

(3) 选择标引方式