

普通高等院校规划教材

文献信息检索理论、 方法和案例分析

何晓萍 主编



机械工业出版社
CHINA MACHINE PRESS

014059543

G254.9

11

普通高等院校规划教材

文献信息检索理论、 方法和案例分析

主 编 何晓萍

副主编 胡翀慧 张 彧

参 编 梁红妮 魏元珍 胡德华

主 审 欧阳伟



机械工业出版社



北航

C1746649

G254.9/11

014020243

本书从理论上系统阐述了数字信息资源检索原理、信息检索方法和信息检索技术；根据信息资源有效使用的时效性，叙述了常用资源平台检索工具的使用方法和技巧；从多元化服务功能着手，结合具体案例，满足读者索取信息、获得知识的个性化需求。

分析、比较和判断检索收集到的大量文献信息，快速锁定特定学科领域具有高影响力的论文；以二次文献作为入口，满足学科整体的需求；以高质量的一次文献，满足纵深研究的需要。可见，科学获取、合理利用文献信息，是开展教学与科学研究的关键。

本书适合作为普通高等院校教学用书，也适合工程技术人员作为参考用书。

图书在版编目（CIP）数据

文献信息检索理论、方法和案例分析/何晓萍主编. —北京：机械工业出版社，2014.9

普通高等院校规划教材

ISBN 978-7-111-47594-1

I. ①文… II. ①何… 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（2014）第 180890 号

机械工业出版社（北京市百万庄大街 22 号 邮政编码 100037）

策划编辑：王海峰 责任编辑：王海峰

版式设计：赵颖喆 责任校对：赵蕊

封面设计：路恩中 责任印制：乔宇

北京铭成印刷有限公司印刷

2014 年 9 月第 1 版第 1 次印刷

184mm × 260mm · 19.5 印张 · 470 千字

0001—3000 册

标准书号：ISBN 978-7-111-47594-1

定价：38.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

电话服务

网络服务

社服务中心：(010) 88361066 教材网：<http://www.cmpedu.com>

销售一部：(010) 68326294 机工官网：<http://www.cmpbook.com>

销售二部：(010) 88379649 机工官博：<http://weibo.com/cmp1952>

读者购书热线：(010) 88379203 封面无防伪标均为盗版

海城王良平時調顯中尉，平沐思學附前千累，**前言** 吳朱姓索餅息前已嚴責惠前劉國

本书编写人员负责的章节为：何晓萍（第1章、第2章）；张或（第3章3.1~3.6节）；胡舸慧（第4章、第5章、第7章）；魏元珍（第3章3.7节，第6章6.2、6.3节）。

网络信息资源与信息检索技术发展迅猛,限于编者的学识水平,书中疏漏和不足之处在所难免,敬请学术界同仁和读者批评指正。

目 录

前言	1	2.3.1 常规法	22
第1章 信息检索基础知识	1	2.3.2 引文法	23
1.1 信息、知识与文献	1	2.3.3 综合法	23
1.2 文献信息类型	1	2.4 文献信息检索步骤	23
1.2.1 按文献信息的物质载体形式划分	1	2.4.1 分析检索课题, 确定检索需求	23
1.2.2 按文献信息的加工程度划分	2	2.4.2 选择文献检索工具, 确认检索字段	24
1.2.3 按文献信息的表现形式划分	2	2.4.3 确定检索词, 编制检索式	24
1.2.4 按文献信息的著录方式划分	5	2.4.4 实施课题检索, 调整检索策略	26
1.3 信息检索途径	6	2.4.5 输出检索结果, 获取所需信息	27
1.3.1 著者途径	6	2.5 检索效果评价	27
1.3.2 题名途径	6	2.5.1 信息检索效果评价的参数	27
1.3.3 号码途径	6	2.5.2 调整查全率和查准率的措施	28
1.3.4 引文途径	7	2.6 检索结果著录格式	28
1.3.5 分类途径	7	2.6.1 著录中的文献类别代码	28
1.3.6 主题途径	7	2.6.2 常用文献著录格式	29
1.4 信息检索语言	7	第3章 网络信息检索技术	31
1.4.1 检索语言的类型	7	3.1 搜索引擎	31
1.4.2 体系分类语言及词表	8	3.1.1 搜索引擎概述	31
1.4.3 受控语言和自然语言	11	3.1.2 搜索引擎工作原理	31
1.5 常用检索技术及其实现	12	3.1.3 搜索引擎的分类	32
1.5.1 布尔逻辑组配检索	12	3.1.4 搜索引擎使用指南	33
1.5.2 截词检索	14	3.1.5 常用搜索引擎	34
1.5.3 位置检索	15	3.2 Google 学术搜索	41
1.5.4 字段限定检索	15	3.2.1 学术搜索概述	41
第2章 计算机文献信息检索	19	3.2.2 Google 学术搜索使用指南	42
2.1 计算机检索系统	19	3.3 网络百科	45
2.1.1 检索系统组织	19	3.3.1 Wiki	45
2.1.2 检索系统功能	19	3.3.2 维基百科 (Wikipedia)	46
2.2 计算机文献信息检索	21	3.3.3 其他常用百科	48
2.2.1 计算机文献信息检索原理	21	3.4 文档分享	51
2.2.2 计算机文献检索形式	21		
2.2.3 数据库检索方式	21		
2.3 文献信息检索方法	22		

3.4.1 文档分享概述	51	4.7.2 数据库检索指南	119
3.4.2 常用文档分享平台	51	4.7.3 注册与个性化服务	123
3.5 在线翻译	54	4.8 Web of Science (SCI、SSCI)	123
3.5.1 在线翻译概述	54	4.8.1 数据库简介	123
3.5.2 常用的在线翻译工具	55	4.8.2 数据库检索指南	124
3.6 移动图书馆	58	4.8.3 注册与个性化服务	133
3.6.1 移动图书馆概述	58	4.9 Ei CompendexWeb (工程索引 数据库)	134
3.6.2 移动图书馆使用指南	59	4.9.1 数据库简介	134
3.7 文献管理软件	65	4.9.2 数据库检索指南	134
3.7.1 文献管理软件定义及主要 功能	65	4.9.3 注册与个性化服务	141
3.7.2 EndNote	65	4.10 SciFinder 网页版 (CA)	141
3.7.3 NoteExpress	69	4.10.1 数据库简介	141
3.7.4 EndNote 和 NoteExpress 比较	71	4.10.2 数据库检索指南	142
第4章 期刊文献	72	第5章 学位论文和会议论文	147
4.1 中国学术期刊网络出版总库 (CAJD)	72	5.1 ProQuest 学位论文全文数据库	147
4.1.1 数据库简介	72	5.1.1 数据库简介	147
4.1.2 数据库检索指南	72	5.1.2 数据库检索指南	147
4.2 万方数字化期刊	85	5.1.3 注册与个性化服务	152
4.2.1 数据库简介	85	5.2 万方博硕学位论文库	152
4.2.2 数据库检索指南	85	5.2.1 数据库简介	152
4.3 维普中文科技期刊数据库 (CSTJ)	91	5.2.2 数据库检索指南	152
4.3.1 数据库简介	91	5.3 中国博硕学位论文全文数据库 (CDFD)	156
4.3.2 数据库检索指南	91	5.3.1 数据库简介	156
4.4 人大复印报刊资料全文数据库	98	5.3.2 数据库检索指南	157
4.4.1 数据库简介	98	5.4 科技会议录索引 (ISTP)	161
4.4.2 数据库检索指南	98	5.4.1 数据库简介	161
4.5 Elsevier_ScienceDirect	105	5.4.2 CPCI-S 数据库检索	162
4.5.1 数据库简介	105	5.5 中国重要会议论文全文数据库 (CPFD)	165
4.5.2 数据库检索指南	105	5.5.1 数据库简介	165
4.5.3 注册与个性化服务	110	5.5.2 数据库检索指南	166
4.6 Springer 电子期刊数据库	111	第6章 专利与标准文献	175
4.6.1 数据库简介	111	6.1 专利文献及其检索	175
4.6.2 数据库检索指南	111	6.1.1 专利基础知识	175
4.6.3 注册与个性化服务	117	6.1.2 国际专利分类表	177
4.7 EBSCO 期刊库	118	6.2 Derwent Innovations Index (德温特专利数据库)	180
4.7.1 数据库简介	118		

6.2.1 德温特专利数据库概况及特点	180	8.1.1 ProQuest Dialog 概况	229
6.2.2 德温特专利数据库记录结构	181	8.1.2 ProQuest Dialog 检索	233
6.2.3 《专利合作条约》	182	8.1.3 ProQuest Dialog 案例分析	239
6.2.4 德温特专利数据库检索途径	183	8.2 读秀学术搜索	242
6.3 网络专利文献	187	8.2.1 数据库简介	242
6.3.1 中国国家知识产权局	187	8.2.2 学术搜索功能	242
6.3.2 中国知识产权网	189	8.3 馆藏书目检索系统	247
6.3.3 万方数据专利数据库	191	8.3.1 大学图书馆馆藏书目检索系统	248
6.3.4 USPTO Patents 美国专利和商标局全文专利数据库	192	8.3.2 中国国家图书馆联机公共目录查询系统	249
6.3.5 esp@cenet 欧洲专利文献数据库	194	8.3.3 CALIS 联合目录公共检索系统	252
6.3.6 世界知识产权组织网站数据库	195	第9章 信息开发与利用	254
6.4 标准文献	198	9.1 定题服务	254
6.4.1 概况	198	9.1.1 定题服务的概念及意义	254
6.4.2 国际标准文献	199	9.1.2 开展定题服务的原则及方法	255
6.4.3 中国标准文献及其检索	200	9.1.3 个性化信息服务和 RSS 技术在定题服务中的应用	257
第7章 电子图书	204	9.1.4 NoteExpress 管理软件在定题服务中的应用	257
7.1 Springer 电子图书	204	9.2 学科服务	261
7.1.1 数据库简介	204	9.2.1 学科服务与 IC 服务	261
7.1.2 数据库检索指南	204	9.2.2 服务模式	261
7.2 EBSCO 电子图书 (原 NetLibrary)	208	9.2.3 学科服务平台	263
7.2.1 数据库简介	208	9.3 科研产出	263
7.2.2 数据库检索指南	208	9.3.1 科研成果概述	263
7.3 超星电子图书	213	9.3.2 科研产出能力	264
7.3.1 数据库简介	213	9.4 科技查新	265
7.3.2 数据库检索指南	214	9.4.1 基础知识	265
7.4 WISEBOOK 电子图书	217	9.4.2 科技查新案例分析	266
7.4.1 数据库简介	217	9.5 科研评价的方法和工具	268
7.4.2 数据库检索指南	218	9.5.1 科研评价的内容	268
7.5 不列颠百科全书在线英文学术版	222	9.5.2 科研评价的功能	268
7.5.1 数据库简介	222	9.5.3 科研评价指标与方法	269
7.5.2 数据库检索指南	222	9.5.4 科研评价的工具	272
7.5.3 个人文件夹	227	第10章 合格论文的撰写	276
7.5.4 其他	228	10.1 学术论文的撰写	276
第8章 其他	229	10.1.1 学术论文的概念	276
8.1 ProQuest Dialog (PQD)	229	10.1.2 学术论文的分类	276

[illegible]

10.2	学位论文的撰写	295
10.2.1	学位论文的概念和分类	295
10.2.2	学位论文的选题、开题	296
10.2.3	学位论文的结构	298
参考文献		301

第1章 信息检索基础知识

1.1 信息、知识与文献

顾名思义,信息是音信、消息的意思。而本质上,信息是客观事物的存在方式和运动状态的反映。中国大百科全书定义,信息是关于事物运动的状态和规律的表征,也是关于事物运动的知识。它用符号、信号或消息所包含的内容,来消除对客观事物认识的不确定性。

知识是人类认识的成果和结晶,是人类在认识和改造世界的社会实践中获得的对事物本质的认识,即指人们对某个事物的熟悉程度。分为陈述性知识、程序性知识、显性知识和隐性知识。它可能包括事实、信息、描述或在教育和实践中获得的技能,也可能是关于理论的或关于实践的。

文献是记录有知识和信息的一切载体,即将知识和信息用文字、符号、图像、音频、视频等记录在一定的物质载体上的结合体。文献由四个基本要素组成:所记录的知识 and 信息(即文献的内容);用于记录知识和信息的物质载体;记录知识和信息的符号;记录的方式或手段。

1.2 文献信息类型

文献信息的类型可以按文献信息的物质载体形式划分、按文献信息的加工程度划分、按文献信息的表现形式划分、按文献信息的著录方式划分等,来揭示不同文献信息源的特点与表现形式。

1.2.1 按文献信息的物质载体形式划分

按文献信息的物质载体形式(即按信息存储载体的物质形态)划分,有印刷型文献、缩微型文献、声像型文献、电子型文献。

1. 印刷型文献

以纸张为存储载体的印刷型文献,是一种历史悠久的传统文献形式,是文献信息传递的主要载体。其优点是便于阅读和流通,不需要特殊设备,传递知识方便、灵活;缺点是存储密度低,占用储藏空间多,不易管理和保存等。

2. 缩微型文献

缩微型文献是以感光材料为载体,以照相为记录手段而形成的一种文献形式,如缩微胶卷、缩微平片、缩微卡片等。缩微型文献的优点是体积小、信息密度高、轻便、易于传递、容易保存。但阅读需要有较复杂的阅读设备来支持,目前较少使用。

3. 声像型文献

声像型文献是通过特定设备,使用声、光、磁、电等技术将信息转换为声音、图像、影

视和动画等形式,给人以直观形象感受的知识载体,也称为视听资料,如唱片、录音带、录像带、CD、VCD、DVD等。声像型文献提供的形象、声音逼真,宜于记载难以用文字表达和描绘的形象资料和音频资料。

4. 电子型文献

电子型文献是以计算机处理技术为核心,把原有的论文、照片、录音和图像等非数字信息,运用数字化技术处理后存储在一定的介质上,成为计算机可以读取、检索的数字信息。有电子图书、电子杂志、电子报纸、联机数据库、光盘数据库、网络数据库等。具有存储密度高、存储速度快、信息处理方便等特点。

1.2.2 按文献信息的加工程度划分

按文献信息的加工程度(即按信息的加工深度和结构等级)划分,可分为一次文献、二次文献、三次文献。

1. 一次文献

一次文献通常是指原始文献,即以作者的研究成果为基本素材而创作(或撰写)的原始制作。包括期刊论文、学术论文、学位论文、科技报告、会议论文、专利说明书、技术标准等公开发表的文献。具有内容新颖丰富,叙述具体详尽,参考价值大等特点。一次文献是最基本的信息源,是文献信息检索和利用的主要对象。

2. 二次文献

二次文献是按照特定目的,对一定范围或学科领域的一次文献进行鉴别、筛选、分析、归纳和加工整理后所得产物,是便于管理和利用一次文献的工具性文献。一般包括目录、题录、索引、文摘等,它以不同的深度揭示单篇文献的外部特征和内容特征,为查找一次文献提供线索,帮助人们在较少时间内获得较多的文献信息。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。

3. 三次文献

三次文献是利用二次文献提供的线索,选用一次文献的内容,经过综合、分析和评述后形成的参考性文献。可分为综述研究类和参考工具类两种类型。前者如动态综述、学科总结、专题述评、进展报告等;后者如年鉴、手册、大全、词典、百科全书等。三次文献源具有系统性、综合性、知识性和概括性等特点。三次文献是一次文献内容的高度浓缩,是研究以往文献内容的重要信息源。

1.2.3 按文献信息的表现形式划分

根据文献信息的表现形式不同,文献可分为以下类型:图书、期刊、会议文献、学位论文、报纸、科技报告、专利文献、标准文献、政府出版物、档案文献和产品资料。

1. 图书

图书是论述或介绍某一学科或领域知识的出版物。主要包括学术专著、参考工具书、教科书等。联合国教科文组织规定,50页以上的印刷品称为图书。图书所论述的内容一般比较系统、全面、成熟、可靠,参考图书资料有助于人们对范围较广的问题获得一般的知识,或对陌生的问题获得初步的了解。但图书编辑出版的周期较长,内容更新慢。

2. 期刊

期刊又称连续出版物,正式刊物持有国家新闻出版总署批准的国内统一连续出版物号。《中国图书馆图书分类法·期刊分类表》将期刊分为五个基本部类,在基本部类中,又分为若干大类。期刊按内容可分为一般期刊、学术期刊、行业期刊、检索期刊四大类。按学术地位可分为核心期刊和非核心期刊两大类。核心期刊是指在某一学科领域(或若干领域)中最能反映该学科的学术水平,信息量大,利用率高,受到普遍重视的权威性期刊。国内主要运用文献计量学的方法,以及通过专家咨询等途径进行核心期刊的测定。非核心期刊经过努力,可以跻身于核心期刊之列;核心期刊如故步自封,也会被淘汰。

期刊出版周期短,内容较新颖丰富,学术性强,能迅速反映国内外的各种学科专业的水平和动向,是人们寻找研究上的新发现、新思想、新见解、新问题的首要信息源。

3. 会议文献

会议文献是国内外学术团体举行的专业会议上发表的论文与报告。会议文献的特点是传递情报比较及时,内容新颖,专业性和针对性强,种类繁多,出版形式多样。它是科技文献的重要组成部分,一般是经过挑选的,质量较高,能及时反映科学技术中的新发现、新成果、新成就以及学科发展趋向,是一种重要的情报源。

会议文献可分为会前文献、会中文献和会后文献三种。①会前文献包括征文启事、会议通知书、会议日程表、预印本和会前论文摘要等。②会议期间的会议文献有开幕词、讲话或报告、讨论记录、会议决议和闭幕词等。③会后文献有会议录、汇编、论文集、报告、学术讨论会报告、会议专刊等。其中会议录是会将论文、报告及讨论记录整理汇编而公开出版或发表的文献。

为更好地利用会议文献,一些国家出版有各种会议文献检索工具或建立数据库,如美国出版的《科技会议录索引》,中国出版的《国内学术会议文献通报》及其数据库等。

4. 学位论文

学位论文是高等院校的毕业生和科研机构的研究生为获得各级学位所撰写的论文,通常采用学士、硕士和博士三级学位制。学士论文是攻读学士学位的本科生所撰写的学位论文,应具有专门知识和技能,具有从事科学技术研究或担负专门技术工作的初步能力;硕士论文是攻读硕士学位的研究生所撰写的学位论文,应在导师指导下,研究生本人独立完成,论文具有自己的新见解,有一定的工作量;博士论文是攻读博士学位的研究生所撰写的学位论文,论题前沿、研究专深,见解独到,是一本自成体系的著作。

由于学位论文具有选题新颖,理论性、系统性较强,阐述详细;参考文献多、全面,有助于对相关文献进行追踪检索的特点,读者对电子论文全文的需求呈上升趋势。目前,常用的学位论文数据库有,中文:清华学位论文全文数据库(CNKI)、万方学位论文全文数据库、中美百万册数字图书项目、国家科技图书文献中心的学位论文库;英文:PQDT博硕士论文数据库。

5. 报纸

报纸是以刊载新闻和时事评论为主的定期向公众发行的印刷出版物。作为大众传播的重要载体,具有反映和引导社会舆论的功能。关于报纸的职能,从不同角度,会得出不同的看法。法国新闻学者贝尔纳·瓦耶纳关于报纸职能的概括,可以被各方面接受:主要的报道职能,随之而来的辩论职能(即传播观点的职能),附带的娱乐职能。

由于报纸所特有的时事性、普及性、大众型,且和社会生活息息相关,所以它拥有最广大的读者。而且不受时域限制、便于阅读,对于阅读或理解能力较低的人,也可相应多耗时间,吸收报纸文章的信息。其缺点是知识不系统,学习分布零乱。即使是网上版报纸的传阅力较传统印刷品报纸强,但受截稿及出版因素影响,不能提供最新资讯以及即时更正信息。

6. 科技报告

科技报告又称研究报告、报告文献。是在科研活动的各个阶段,由科技人员按照有关规定和格式撰写的,以积累、传播和交流为目的,能完整而真实地反映其所从事科研活动的技术内容和经验的特种文献。每份报告自成一册,通常载有主持单位、报告撰写者、密级、报告号、研究项目号和合同号等。按内容可分为报告书、论文、通报、札记、技术译文、备忘录、特种出版物。大多与政府的研究活动、国防及尖端科技领域有关,发表及时,课题专深,内容新颖、成熟,数据完整,且注重报道进行中的科研工作,是一种重要的信息源。

7. 专利文献

专利文献是专利申请人向专利局递交的有关发明目的、构成和效果的技术性法律文件,经专利局审核后,各国专利局的正式出版物。如专利说明书、专利公报、专利文摘、专利索引、专利分类表等。

专利文献具有技术性、新颖性、独创性、实用性、具有法律效力等特征。其作用是:对专利申请进行专利性检索;启迪发明创造思路;可以了解该领域的最新动态;有利于技术转让;有利于企业的技术开发;有利于引进国外先进技术和设备;作为专利诉讼的有力依据;是重要的技术经济情报来源。

目前,中国专利检索系统有国家知识产权局专利检索系统、中国专利信息中心专利检索系统、万方专利技术数据库等。欧洲专利检索系统有 esp@cenet 网站、世界专利数据库、EP 专利数据库。美国专利检索系统有美国专利商标局专利检索系统、授权专利 (Issued Patent) 检索、申请公布 (Published Application) 检索。

8. 标准文献

标准文献是技术标准、技术规范和技术规则等文献的总称,具有法律效应和时效性,约束力强,是科研生产活动的重要依据和信息来源。一个国家的标准文献反映该国的生产工艺水平和技术经济政策,而国际现行标准则代表了当前世界水平。各国标准有效期不同,ISO 国际标准化组织的标准每 5 年复审一次,平均标龄为 4.92 年。我国国家标准有效期一般为 5 年。

9. 政府出版物

政府出版物是政府用以发布政令和体现其思想、意志、行为的物质载体,同时也是政府的思想、意志、行为产生社会效应的主要传播媒介。由政府机构制作出版或由政府编辑、由政府指定出版商出版的文献。目前各国对政府出版物尚无一致定义。大致可分为两类:一类是行政性文件,包括会议记录、司法资料、条约、决议、规章制度以及调查统计资料等;另一类是科技性文献,包括研究报告、科普资料、技术政策文件等。

政府出版物数量巨大,内容广泛,出版迅速,资料可靠,是重要的信息源。由于政府出版物与其他类型文献有一定重复,如图书、报告、会议录等多种形式,因此,可根据具体情况按具体出版形式加以识别。

10. 档案文献和产品资料

国家和地方各级机关、企事业单位、社会团体等在进行日常工作时所产生的, 并经立卷、归档的各种资料, 如电文、会议记录、报表、文件原稿、人事材料等。档案一般为内部使用, 不公开发行, 有些有密级限制, 因此在参考文献和检索工具中极少使用。

产品资料一般是指产品样本, 即产品说明书。好的产品说明书含有丰富的内容, 包括产品规格、产品特点、产品专利号等多种对生产有用的信息。

1.2.4 按文献信息的著录方式划分

根据用户对文献局部信息需求的不同, 记录文献信息知识的著录格式也不同。通常, 按文献信息的著录方式划分, 可分为目录、题录、文摘、索引。

1. 目录

目录又称书目, 是以一个完整出版物(一本书、一种报纸或一本期刊等)为单位, 对出版物外表特征进行著录, 如出版物名称、著者、出版项(出版者、出版地等), 没有具体内容。目录作为一种联系文献与需求者之间的媒介或纽带, 以最大限度满足人们的书目情报需求为目的对文献信息进行科学的揭示和有效的报道, 并且不受时间和空间的限制。如图书目次是揭示和报道图书的工具, 馆藏目录是揭示图书馆各自收藏书刊资料的目录。图书馆中最常用的有分类目录、书名目录、著者目录和主题目录。联合目录是汇总若干图书馆或其他收藏单位收藏的书刊资料编制的目录。可了解馆际之间图书引荐情况, 达到资源共享、互通有无。

2. 题录

题录是用来描述某一文献的外部特征并由一组著录项目构成的一条文献记录, 利用它可以准确地鉴别一种出版社或其中的一部分。题录通常以一个内容上独立的文献单元(如一篇文献, 一本书)为基本著录单位。著录项目包括题名、著者、出处等外部特征, 不揭示文献内容特征。将一系列题录有序排列, 即构成“目录”或“文献通报”。借助于题录可查找最新的相关文献线索, 如题名中与主题相符的相关文献; 著者项中是学科带头人发表的相关文献; 从出处中筛选具有学术权威刊物的相关文献。

3. 文摘

文摘是检索刊物中描述文献内容特征和外部特征的条目, 包括文献摘要和题录部分的题名、著者、出处等, 是对文献内容作实质性描述的文献条目。有报道性、指示性、报道-指示性三种文摘类型。文摘条目及文摘杂志具有报道、检索、参考和交流等功能, 是开展信息交流的重要手段。其主要作用有: ①通报最新科学文献; ②节省阅读时间; ③引导检索原文; ④能够获取因语言障碍无法得到的文献信息。

4. 索引

索引是将图书、期刊等文献中的论文按内容进行分类标引或主题标引后, 严密组织编排, 以题录的形式注明论文出处。它是以一种完整出版物中的某一知识单元为著录对象, 每一条著录项就是一条题录(可混称索引为题录)。

索引还有另一个含义, 指某种检索途径, 如主题索引、著者索引、分子式索引等索引是揭示文献内容出处, 提供文献查考线索的工具书。按照索引款目的标目, 可分为题名索引、著者索引、语词索引、主题索引、分类索引、引文索引、文献序号索引和代码索引等。按照

索引收录文献的类别,可分为书后索引、期刊索引、报纸索引、专利索引、标准索引等。按照索引的载体形式可分为卡片式索引、附录式索引、单卷式索引和期刊式索引(即索引期刊)等。

所有索引款目实现有序化编排。将文献中的一些检索标识(如人名、题名、分类号、主题词等)分别摘录出来,并注明它们所在的位置(如题录号、文摘号等),再按一定的规则排列和组织,一般只起指引特定信息内容及其存储地址的作用。其本质特征是只揭示内容出处或文献线索,并不直接提供事实、资料本身。

数据库中,索引是一个单独的、物理的数据库结构,它是某个表中一列或若干列值的集合和相应的指向表中物理标识这些值的数据页的逻辑指针清单。数据库使用索引的方式与使用书籍中的索引的方式很相似,搜索索引以找到特定值,然后顺指针找到包含该值的行。主要功能是为人们准确、迅速地获得文献资料提供线索性指引。常见的索引主要有报刊论文资料索引、文集篇目索引、语词索引、文句索引、关键词索引、专名索引、主题索引等。

1.3 信息检索途径

根据文献信息资源的外部特征和内部特征,可确定文献信息检索途径,即检索点或检索入口。文献信息资源的外部特征一般由作者项、题名项、出处项组成。责任者(著者、译者、编者等)形成著者途径;题名(书名、刊名等)形成题名途径;ISSN号、ISBN号、专利号等形成号码途径;引用文献形成引文途径。文献信息资源的内容特征主要有学科属性形成的分类途径和分析构成文献信息内容要素的主题途径。而主题途径中运用较广的有主题词途径和关键词途径两种。常用的文献信息检索途径有著者途径、题名途径、号码途径、引文途径、分类途径、主题途径等。

1.3.1 著者途径

利用检索系统中著者、编者、译者、专利权人、机关团体及机构名称等字段进行检索的途径统称为著者途径。从著者途径入口,查找已知著者、译者、编者的学科代表作,可以系统地发现和掌握该学科的研究和发展,了解某一作者的最新论著。同时,通过对作者的文献信息检索结果,可以链接相同学科的作者,进一步深入、系统地浏览、搜索相关文献信息。

1.3.2 题名途径

检索系统提供按题名字段检索的途径。通过书名、刊名、文献篇名等题名途径入口,模糊或精确地查找与题名的主题相关的文献信息。同时,也可以链接检索结果中的参考文献题名,从中浏览筛选相关文献信息。

1.3.3 号码途径

从号码途径入口,查找已知专用号码和专用符号代码,直接或间接(利用代码索引)检索相关文献信息。专用号码如报告号、专利号、合同号、标准号、国际标准书、刊号(ISBN号、ISSN号)等;专用符号代码如元素符号、分子式、结构式等。

1.3.4 引文途径

文献所附参考文献或引用文献,是文献的外表特征之一。利用这种引文而编制的索引系统,称为引文索引系统。提供从被引论文去检索引用论文的一种途径,称为引文途径。

由于文献之间引用与被引用的关系可视为文献内容之间具有联系的特性。换言之,两篇文献发生引用关系,那么它们在主题上就是相关的。引文途径检索,通过对一般文献最后部分附有的参考文献或引用文献链接的追溯查找,或利用引文索引系统的文献检索,其结果更具有科学性和拓展性。

1.3.5 分类途径

分类途径是以知识体系为中心按学科分类体系来编排检索文献的途径,因此,比较能体现学科系统性,反映学科与事物的隶属、派生与平行的关系,有利于从学科所属范围来查找文献资料,并且可以起到“触类旁通”的作用,是提高文献检索查全率的有效途径。但不便于交叉学科或科目名称属性不明的文献信息检索。

从分类途径检索文献资料,主要是利用分类目录和分类索引。目前,我国通用的分类法主要有《中国图书馆分类法》,简称《中图法》;《中国科学院图书馆图书分类法》,简称《科图法》。其中《中图法》是国家推荐统一使用的分类法,使用范围最广泛。

1.3.6 主题途径

主题途径是指用户通过反映文献资料内容的主题词或关键词来检索文献。由于主题法能集中反映一个主题的各方面文献资料,因而便于读者对某一问题、某一事物和对象作全面系统的专题性研究。通过主题目录或索引实施检索,即可查到同一主题的各方面文献资料,获得的信息专指性强,达到精确检索的效果,是提高文献检索查准率的有效途径。

但是,主题途径检索的结果查全率较低。而且,主题词一般是以规范化的词汇来表达文献内容的主题,因此,主题途径对用户的检索知识要求较高。

1.4 信息检索语言

检索语言(Retrieval Language)是描述检索系统文献特征并表达用户信息提问的一种专门语言。它能使文献存储者和检索者达到共同理解,实现存取统一。其实质是表达文献主题的一系列概念标识。文献检索包括存储和检索两个部分。存储是指编制检索工具和建立检索系统;检索则是利用这些检索工具和检索系统来查找所需的文献,连接文献存储和检索这两个密切相关过程的正是检索语言。检索语言作为标引和检索之间的约定语言,是沟通文献存储和检索两个过程、标引人员和检索人员双方思想的桥梁。

1.4.1 检索语言的类型

检索语言就其组配方式来划分,可分为先组式检索语言和后组式检索语言;就其表达文献特征来划分,可分为描述文献外表特征的检索语言和描述文献内容特征的检索语言。描述文献外表特征的检索语言,最常用的有引文语言、题名(书名、刊名、文献篇名等)、著者、文献

号码(索书号、ISSN号、专利号、报告号等)。描述文献内容特征的检索语言,包括体系分类语言、关键词语言、单元词语言、标题词语言、叙词语言等。其类型如图1-1所示。



图 1-1 检索语言类型

检索语言的种类虽然很多,但是在常用的检索工具中所使用的检索语言,主要是体系分类语言、受控语言、自然语言。不同的检索语言构成不同的标识和索引系统,以提供不同的检索点和检索途径。

1.4.2 体系分类语言及词表

分类是区别事物及其相互关系的一种思想方法,是人们思维活动中的一种本能。体系分类是一种直接体现知识分类的等级概念的标识系统。每个类目都用分类号做标识,每个分类号都是表达特定知识概念的词汇,即体系分类语言的语词。分类号的作用,表示类目在分类体系中的位置,表示类目的排列顺序。

1. 体系分类法的特点

体系分类法是将文献根据其所属的学科内容分门别类地系统化组织的一种方法。它以学科门类为基础,根据文献的内部和某些外部特征,运用概念划分的原则,按知识门类的逻辑次序由总体到分支、由一般到具体、由简单到复杂进行层层划分,逐级展开。一个大类或上位类每划分一次便产生许多子类目,所有不同级别的子类目向上层层隶属,向下级级派生,从而形成了一个严格有序的、直线性知识门类等级体系。它以学科的衍生、发展为基础进行分类,较好地体现了学科的系统性,反映了事物的隶属、平行、派生关系,有利于人们从专业、学科的角度检索文献。

体系分类法由于可以直接体现知识分类的要求,把众多的文献强制性地按照特定的分类体系予以组织,提供从学科分类检索文献的途径,具有直线性序列特点和较强的系统性,从而为文献的分类组成、编排存储和类目扩展等发挥了独特的功能。故便于检全某一学科、某一专业的文献,又可以根据课题的需要扩大或缩小检索范围。因此,对于系统掌握和利用一个学科或专业范围的知识来说,它是方便和有效的。

但是,体系分类法采用从上位类到下位类逐级展开细分而形成的纵向线性关系,难以反映当代科学技术交叉渗透而出现的多元知识空间,不适于标引和检索主题因素复杂的多维概念问题;再者,由于科学技术日新月异的动态,体系分类法的类目设置总是落后于科学技术