



普通高等教育“十三五”规划教材

101010

文献信息检索

WENXIAN XINXI JIANSUO

申 燕 主编



非
外
借



中国纺织出版社

普通高等教育“十三五”规划教材

文献信息检索

主 编 申 燕
副主编 徐宝凤 孔庆国
参 编 赵 晶 吴 暇 袁春霞
钟盈盈

图书在版编目 (CIP) 数据

文献信息检索 / 申燕主编. -- 北京: 中国纺织出版社, 2016.9

ISBN 978-7-5180-2916-7

I. ①文… II. ①申… III. ①信息检索—高等学校—教材 IV. ①G254.9

中国版本图书馆 CIP 数据核字 (2016) 第 210420 号

策划编辑: 张向红

责任编辑: 张向红

责任设计: 林昕瑶

责任印制: 储志伟

中国纺织出版社出版发行

地 址: 北京市朝阳区百子湾东里 A407 号楼 邮政编码: 100124

邮购电话: 010-67004461 传真: 010-87155801

<http://www.c-textilep.com>

E-mail: faxing@c-textilep.com

北京市迪鑫印刷厂印刷

各地新华书店经销

2016 年 9 月第 1 次印刷

开 本: 787mm × 1092mm $1/_{16}$ 印张: 14.5

字 数: 330 千字 定价: 36.00 元

凡购本书, 如有缺页, 倒页, 脱页, 由本社图书营销中心调换

前 言

21 世纪是信息化、数字化、全球化的知识经济时代。信息意识和信息检索及利用能力是本世纪人才应具备的一项基本素质。作为新时代的大学生,信息检索能力的培养显得越来越重要。文献检索课已成为我国高等院校培养大学生信息素质的公共课之一。

随着信息科技步入多媒体发展时期,计算机信息检索已进入了应用与发展阶段,以 Internet 为代表的全球性网络的实际应用进一步推动了这一发展,这既是对手工检索的扩展,也是时代进步的需要。

教材是教学内容的主要依据,是实现教学预期目标的工具。而在当前适合应用型本科院校层次的教材很少,我们在吸收众多同类教材优点的基础上,根据课程改革的要求和计算机检索发展的要求编写了本教材。

本教材层次结构清晰,逻辑性强,注重检索过程的实际操作,难易程度适合高职高专及应用型本科院校学生的实际。

本教材在编写过程中,参照和引用了国内外许多专著、论文、教材和网站的有关观点,在此对提供文献资料的作者表示由衷的感谢!

由于编者水平有限,难免存在不足和疏漏之处,敬请同仁及广大读者批评指正。

编 者

目 录

第1章 文献信息检索概述	1
1.1 文献信息基本知识	1
1.1.1 文献及其相关概念	1
1.1.2 文献的属性	3
1.1.3 文献的分类	3
1.2 文献信息检索	6
1.2.1 文献信息检索的涵义、类型和基本原理	6
1.2.2 检索语言及其类型	8
1.2.3 文献信息检索的途径	15
1.2.4 文献信息检索的方法和步骤	17
1.3 文献信息检索工具	21
1.3.1 文献信息检索工具的概念、特点与分类	21
1.3.2 检索工具的常见类型介绍	21
1.3.3 检索工具的结构和刊名缩写问题	25
第2章 网络信息检索	26
2.1 网络信息检索概述	26
2.1.1 网络检索的产生	26
2.1.2 网络检索的特点	26
2.2 网络信息资源	27
2.2.1 网络信息资源的种类	27
2.2.2 网络信息资源的特点	30
2.3 网络检索工具	31
2.3.1 搜索引擎的定义	31
2.3.2 工作原理	31
2.3.3 搜索引擎的基本检索功能	32
2.3.4 搜索引擎的类型	33
2.4 常用中文搜索引擎	34
2.4.1 Google 搜索引擎	34
2.4.2 百度中文搜索引擎	38

2.5 网络资源目录	40
2.5.1 目录型网络检索工具简介	40
2.5.2 常用中文目录型网络检索工具	41
2.5.3 常用英文目录型网络检索工具——Yahoo	43
2.6 多元搜索引擎	44
2.6.1 概述	44
2.6.2 常用多元搜索引擎	45
2.7 教育信息检索——中国高等教育文献保障系统(CALIS)	47
2.7.1 CALIS 简介	47
2.7.2 CALIS 的服务功能	47
2.7.3 CALIS 引进网络数据库简介	48
2.8 学科专业资源特色网站选介	50
2.8.1 中国经济信息网(http://www.cei.gov.cn)	50
2.8.2 北大法律信息网(http://www.chinalawinfo.com)	51
2.8.3 中国工程技术信息网(http://www.cetin.net.cn)	52
2.8.4 当当书店(www.dangdang.com)	53

第3章 中文数据库检索

54

3.1 CNKI 知识资源网站	54
3.1.1 CNKI 资源概况	54
3.1.2 检索页面介绍	55
3.1.3 检索方法	56
3.2 万方数据库	64
3.2.1 概述	64
3.2.2 主要数据库简介	64
3.2.3 检索方法	65
3.2.4 文章浏览与编辑	68
3.3 维普数据库	69
3.3.1 维普数据库概述	69
3.3.2 CQVIP 中文科技期刊全文数据库检索	69
3.3.3 CQVIP 中文科技期刊资源使用(题录/文摘检索)	73
3.3.4 OpenUrl 增值服务	76
3.3.5 辅助功能	76
3.4 中国人民大学报刊复印资料全文数据库	77
3.4.1 概述	77
3.4.2 检索方法	77
3.4.3 查询词帮助功能	79
3.5 电子图书数据库	80
3.5.1 电子图书与电子图书馆	80

3.5.2 超星数字图书馆	80
3.5.3 书生之家电子图书	84
第4章 外文数据库的检索	88
4.1 《工程索引》	88
4.1.1 Engineering Village 2 页面介绍	88
4.1.2 Engineering Village 2 检索方法	88
4.1.3 检索结果的分析与显示	92
4.2 INSPEC 数据库概述	93
4.2.1 INSPEC 数据库概述	93
4.2.2 INSPEC 的特点	94
4.2.3 INSPEC 的学科范围	94
4.2.4 INSPEC 数据库检索	95
4.2.5 检索字段	95
4.2.6 检索方法	97
4.2.7 检索结果处理	99
4.2.8 检索历史及个性化服务	99
4.3 SCI 与 SSCI 数据库	100
4.3.1 概述	100
4.3.2 Web of Science	101
4.4 EBSCO 系列数据库	106
4.4.1 数据库简介	106
4.4.2 选择数据库	107
4.4.3 检索方法	107
4.4.4 检索结果处理	112
4.5 ASME 电子全文期刊数据库	113
4.5.1 ASME 简介	113
4.5.2 Scitation 平台介绍	113
4.5.3 Scitation 平台检索方法	114
4.5.4 ASME 电子期刊数据库的检索	116
4.6 Science Direct 全文电子期刊数据库	117
4.6.1 数据库介绍	117
4.6.2 SDOL 数据库的检索方法	118
4.6.3 检索技术	119
4.6.4 检索结果处理	119
第5章 特种文献的检索	121
5.1 科技报告	121
5.1.1 科技报告的类型	121

5.1.2	国内科技报告及其检索	122
5.1.3	获取科技报告原文的方法	127
5.2	会议文献及其检索	128
5.2.1	会议文献	128
5.2.2	国内会议文献检索工具	129
5.2.3	国外会议文献检索工具	132
5.3	标准文献的检索	138
5.3.1	标准文献基础知识	138
5.3.2	标准文献的分类	139
5.3.3	中国标准文献检索	141
5.4	学位论文及其检索	145
5.5	专利文献及其利用概述	146
5.5.1	专利文献的概念、特点和类型	146
5.5.2	专利文献的编排结构	148
5.5.3	国际专利分类法	149
5.5.4	英国德温特分类系统	152
5.5.5	专利文献的检索与利用所面临的问题	152
5.6	中国专利文献的检索	153
5.6.1	中国专利概况	153
5.6.2	中国专利文献的编排结构	153
5.6.3	中国专利文献的编号体系	156
5.6.4	中国专利文献的计算机检索	159
5.7	外国专利文献的检索	162
5.7.1	德温特印刷型专利检索工具简介	162
5.7.2	《世界专利索引》	163
5.7.3	外国专利文献网上检索系统和数据库	168
第6章	参考工具书	173
6.1	参考工具书的涵义、特征	173
6.1.1	参考工具书的涵义	173
6.1.2	参考工具书的特征	173
6.2	参考工具书的类型及功用	174
6.2.1	字典、词典	174
6.2.2	类书	178
6.2.3	政书	179
6.2.4	百科全书	180
6.2.5	年鉴	182
6.2.6	手册	184
6.2.7	资料汇编	185

6.2.8 表谱	187
6.2.9 图录	190
6.2.10 名录	191
第7章 文献合理利用与论文写作	193
7.1 学术规范	193
7.1.1 学术道德规范	193
7.1.2 学术法律规范	193
7.1.3 学术引文规范	194
7.1.4 写作技术规范	195
7.2 文献的合理使用	195
7.2.1 合理使用概述	195
7.2.2 传统文献的合理使用	196
7.2.3 电子文献的合理使用	197
7.2.4 学术造假与剽窃	198
7.3 学术论文的撰写	200
7.3.1 学术论文概述	200
7.3.2 学术论文的编写格式	200
7.3.3 数字的使用规则	208
7.3.4 图表的设计和制作原则	210
7.3.5 中图分类号和文献标识码的选取	210
7.4 学术论文的投稿	211
7.4.1 国际学术成果发表制度	211
7.4.2 投稿与审稿结果的处理	212
7.4.3 学术论文成功发表的策略	215
7.4.4 二次发表与一稿多投	216
7.4.5 常用投稿网址	217

第1章 文献信息检索概述

现代社会已进入信息时代。通信技术、计算机技术和声像技术等现代信息技术的发展及相互融合,拓宽了信息的传递和应用范围,使人们在广阔范围内可随时随地获取和交换信息。信息与物质、能量构成了现代社会的三大资源,成为社会进步的强劲推动力。在信息社会,信息量急剧增长,信息载体呈多样化,如何有效地检索、处理和利用信息资源,已成为人们亟待解决的问题。

我们所处的时代是一个前所未有的不断产生、传递和利用信息情报的时代。因此,有人称我们的时代为信息时代。那么,究竟什么是信息、什么是知识,它们与本书大量讨论的文献有着什么关系,这是我们首先需要了解的问题。

1.1 文献信息基本知识

1.1.1 文献及其相关概念

1. 文献

文献的定义多种多样。根据《中华人民共和国国家标准:文献著录总则》(GB3792.1-83),文献是记录有知识的一切载体。国际标准化组织《文献情报术语国际标准》(ISO/DIS5217)则将文献解释为:“在存储、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储有信息或数据的载体。”也有学者认为“文献是用符号、声像等记录在一切载体上的知识。”

从上述文献的定义可以看出,文献具有二个基本要素:一是文献含有知识信息;二是负载知识信息的物质载体;三是记录知识信息的符号和技术。文献记录知识信息,而这些知识信息又依附于载体而存在。

文献是记录知识信息的物质形式,也是借以传递知识信息的工具。由于有文献的存在,人类的知识才得以保存和传播,人类的科学技术和文化才得到继承和发展。文献不仅包括书刊等印刷型出版物,而且包括会议文献、科技报告、专利文献、学位论文、科技档案等各种类型的出版物,也包括古代的甲骨文、竹简、帛书等,以及当今的声像出版物、电子出版物和 Internet 网上的信息。

文献不仅指知识,同时还包括载体、信息媒体(信息或知识的表达方式或技术手段)。载体是物理的,信息媒体是逻辑的;信息媒体类似于计算机的操作系统的平台,例如,Unix、Linux、Windows。

2. 信息

什么是信息?半个世纪以来,科学界一直在对其定义进行积极的探讨。信息的定义

在不同的领域人们赋予它不同的定义。关于信息的定义多达数十种,它们都从不同的角度反映了信息的某些特征。到目前为止,尚无一种定义被社会各界一致接受。概括起来主要有以下观点:《辞海》(1989年版)对信息的解释是:①音讯、消息。②通信系统传输和处理的对象,泛指消息和信号的具体内容和意义。

还有些观点认为:信息是消息、情报、信号、数据和知识;信息是通过文字、数据和各种信号来传递、处理和表现客观事物特性的知识流。

对信息具有广泛影响的定义为:信息是指应用文字、数据或信号等形式通过一定的传递和处理,来表现各种相互联系的客观事物在运动中所具有的特征性内容的总称。

总之,信息是对客观世界中各种事物的变化和特征的反映;是客观事物之间相互作用和联系的表征;是客观事物经过感知或认识后的再现。信息普遍存在于整个宇宙之中,信息无处不在、无时不有,是人们认识世界、改造世界、取之不尽、用之不竭的宝贵资源。信息的增长速度和利用程度已成为现代社会文明和科技进步的重要标志之一。

3. 知识

知识是信息的一部分,是人们对客观事物存在和运动规律的认识。它提供某种经过思考的判断和某种实验的结果。

《辞海》对知识的解释为:人类认识的成果或结晶。包括经验知识和理论知识。经验知识是知识的初级形态,系统的科学理论是知识的高级形态。人的知识是后天在社会实践中形成的,是对现实的反映。

知识借助于一定的语言形式,或物化成为某种劳动产品的形式,可以交流和传递给下一代,成为人类共同的精神财富。知识随社会实践、科学技术的发展而发展,一般可以分成三大类:自然科学知识、社会科学知识和思维科学知识。哲学知识则是关于自然、社会和思维知识的概括和总结。

换言之,知识是人类在改造客观世界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加工的过程,而认识飞跃的结果即为知识。也就是说信息被有选择地收入到人们的思维系统,经过人们大脑的储存、识别、加工、处理、转换等而形成知识。

人们不仅能通过信息感知世界、认识世界和改造世界,而且能将获得的信息转变成知识作为认识和改造世界的武器。把信息转化为知识,再把知识转化为智慧,是一种动态过程,是一种开拓过程。反过来,智慧又会转化为新知识,新知识又会转化为新信息,人们通过一定的手段和社会传递过程,借助媒体传给使用者。

4. 文献、信息、知识之间的关系

信息的内涵和外延在不断扩大,并渗透到人类社会和科学技术的众多领域,人类在接受了来自人类社会及自然界的大量信息,通过认识、分析和重新组合,使信息系统化而形成知识。知识是人类大脑加工提炼信息的成果,是同类信息的深化、积累,所产生的新的知识又会转化为新的信息,如此循环反复。知识依附于载体上就是文献。文献是传递知识信息的介质,是固化了的知识信息。

综上所述,系统化的信息就是知识,知识是信息的一部分,文献是知识的一种载体。文献不仅是知识传递的主要物质形式,也是吸收利用信息的主要手段。

1.1.2 文献的属性

文献本身所固有的性质包括四个方面:

一是知识信息性。这是文献的本质属性。任何文献都记录或传递一定的信息知识。离开知识信息,文献便不复存在。传递信息、记录知识是文献的基本功能,人类的知识财富正是依靠文献才得以保存和传播的。

二是客观物质性。文献所表达的知识信息内容必须借助一定的信息符号、依附于一定的物质载体,才能长时期保存和传递。

三是人工记录性。文献所蕴涵的知识信息是通过人们用各种方式将其记录在载体上的,而不是天然荷载于物质实体上的。

四是动态发展性。文献并非处于静止状态,而是按新陈代谢规律运动着。随着人类记录水平的提高,信息交流的频繁,文献的数量日益庞大,形式日益多样;与此同时,文献的更新速度也在加快,生命周期日益缩短,形成了有规律的运动。

1.1.3 文献的分类

根据不同的划分标准,文献可分成多种类型。分类的标准主要有记载内容、记录方式、载体材料等等。

1. 按载体形式划分

(1) 印刷型文献:是以纸张为载体,以印刷(油印、胶印、铅印、影印、复印等)为记录手段而产生的具有悠久历史的一种文献形式,是迄今为止仍占主导地位的一大文献类型。其优点是用途较广、读取方便、流传不受时空限制等。其缺点是存储密度低、较笨重、占据空间大。

(2) 缩微型文献:是以感光材料为载体,以缩微照相为记录手段而产生的一种文献类型,包括缩微平片和缩微胶卷等。优点是:体积小、价格低、存储信息密度高,便于收藏、保存和传递。其缺点是必须借助缩微阅读机,使用不方便。

(3) 声像型文献:是以磁性材料或感光材料为载体,以磁记录或光学技术为手段直接记录声音、视频图像而形成的一种文献。如唱片、录音录像带、幻灯片、电影等。其优点是:生动直观;缺点是成本较高且不易检索和更新。

(4) 电子文献:是近年来由于计算机和网络的广泛应用而产生的一种新文献。电子文献系指以数字代码方式将图、文、声、像等信息存储在磁光电介质上,通过计算机或具有类似功能的设备阅读使用,用以表达思想、普及知识和积累文化的文献。它包括正式出版的电子文献(即电子出版物,如电子图书、电子期刊、电子报纸等),也包括非正式出版的电子文献(如校园上的各类行政报告、网上的会议资料、内部电子期刊、电子教程、信函等)。

电子文献按贮存内容的表现形式划分有:电子图书,电子期刊,电子报纸,数据库,音像,多媒体,程序,文档;按信息存贮载体的不同划分,电子文献可分为:磁带、磁盘,光盘,集成电路卡(包括各类游戏卡、磁卡等),网络。上述的磁带、磁盘、光盘、集成电路卡等电子文献又可统称为制品型(或封装型)电子文献。

电子文献的存贮、阅读和查找利用都须通过计算机、网络才能进行,所以既有信息量

大、查找迅速的优点,又有设备昂贵、使用费用高的缺点。

2. 按加工层次划分

根据加工层次不同,可将文献划分为零次文献、一次文献、二次文献、三次文献四个等级。

(1) 零次文献:指记录在非正规物理载体上,未经出版发行的或未进入社会交流的最原始的文献。如私人笔记、手稿、考察记录、试验记录、原始统计数字、技术档案等。其主要特点是内容新颖,但不成熟,不公开交流,难以获得。

(2) 一次文献:指以作者本人的生产与科研工作成果为依据而创作的原始文献。如专著、期刊论文、科技报告、会议论文、专利文献、学位论文等。一次文献真实、具体、参考使用价值高,但也有分散、数量庞大而查阅不便的缺点。

(3) 二次文献:也称检索性文献,指对一次文献进行精选、提炼、浓缩和加工,标引出文献的主题,编制成具有多种检索途径的检索工具,如文摘、索引、题录等。二次文献具有浓缩性、汇集性、有序性等特点,它是查找一次文献的工具或手段。

(4) 三次文献:也称参考性文献,是指利用二次文献的线索,系统地检索出一批相关文献,并对其内容进行综合、分析、研究和评述而编写出来的文献。如述评、动态综述、进展报告、数据手册、年鉴等。

三次文献可再分为:①文献型(知识浓缩型):如综述、述评、专著之类;②数据型:如字典、词典、数据手册、百科全书等。其中文献型为三次文献的主要代表。

另外,还有半文献(灰色文献)的说法。半文献是指未公开出版的内部文献,往往是通过正规售书途径得不到的资料,主要通过交换、赠送或其他途径获得。近年出现的私人网站、博客(Blog)可视为半文献。

3. 按出版类型划分

根据出版类型的不同,文献大体可分为如下13种类型。

(1) 图书:是单册出版的正式公开出版物。这是历史最悠久的文献类型。图书的特点是内容一般比较成熟,代表了某一时期某一学科的发展水平,但出版周期较长。

图书有书号,如:ISBN 0-13-165316-4 是书号。其中 ISBN 是国际标准书号,后面共有10位数字,这10位数字分为四段,每段之间用短横线相连。零为第一段组号,是语言区域代码。13为第二段组号,是出版者代号。165316是第三段组号,又叫书序号或书名。这里的165316是指第165317种,因为第一种为零。4为第四段组号,又叫检验码或校验号。它的作用是可用其检查 ISBN 后面的数字是否有误。近年来国际标准化组织(ISO)发布了 ISBN 标准的新版本,对目前使用的1972年 ISBN 标准进行了修正,即自2007年1月1日起,ISBN号将变为13位数,并且首次要与相应的条形码相一致。

(2) 期刊:是指定期或不定期出版的有固定名称的连续出版物。其特点是出版周期短,报道速度快,数量大,内容丰富新颖,能及时反映当代社会和科技的发展水平和动向。

(3) 会议文献:指各种学术会议上发表的论文和报告。会议文献学术性较强,往往反映了当前的学科进展和发展动态,是获取最新信息的重要来源。

(4) 研究报告:是研究单位和个人向上级或委托单位撰写的关于某个课题研究成果的正式报告。其出版特点是各篇单独成册,统一编号,由主管机构连续出版。在内容方面,报告比期刊论文等专深、详尽、可靠,是一种不可多得的情报源。科技报告是科技人

员从事某一专题研究所取得的成果和进展的实际记录。其特点是反映新技术、新学科较快,内容比较专深、新颖,数据比较可靠,保密性较强,有相当一部分科技报告不公开发行。

(5) 专利文献:专利文献是指根据专利法公开的有关发明的文献,主要为专利说明书,也包括专利法律文件和专利检索工具。专利文献具有新颖性、创造性和实用性的特点,且范围广泛、出版迅速、格式规范。

(6) 学位论文:是指高等院校的博士研究生、硕士研究生、本科生毕业时所撰写的学术性研究论文。学位论文(博士论文、硕士论文)具有一定的独创性,论及的问题比较专深、详尽,有较高的参考价值。

(7) 标准文献:是对工农业新产品和工程建设的质量、规格、参数及检验方法所做的技术规定。它是一种经权威机构批准的规章性文献,具有一定的法律约束力。

(8) 政府出版物:是各国政府及所属机构颁布的文件,如政府公报、会议文件和记录、法令汇编等等。所包括的内容范围广泛,几乎涉及整个知识领域,但重点则在政治、经济、法律、军事等方面。政府出版物具有正式性和权威性的特点。

(9) 产品样本:是对定型产品的性能、构造、原理、用途、使用方法和操作方法、产品规格等所作的说明。它往往配有外观照片、结构图,直观性强、技术成熟。

(10) 技术档案:指生产建设、科技部门和企事业单位针对具体的工程或项目形成的技术文件、设计图样、图表、照片、原始记录的原本及复印件。包括任务书、协议书、技术经济指标和审批文件、研究计划、研究方案、试验记录等。它是生产领域、科学实践中用以积累经验、吸取教训和提高质量的重要文献。科技档案具有保密性,常常限定使用范围。

(11) 报纸:是一种报道及时,内容广泛,文字通俗的信息源,其中的广告、新闻报道更是重要的信息源。由于报纸信息累积信息量大且杂乱无章,不借助检索工具查找起来很困难,因此重要的报纸都编有月度或年度索引,有的报纸文章与期刊论文一起被摘录编排,形成报刊索引。

(12) 电子预印本文献(e-preprint):即电子版的学术文献,它是学术论文的数字形式,作为一类重要的免费电子资源,越来越被人们重视。其实质包括预印本(Preprint)与后印本(Postprint)两种形态的文献。预印本指论文原稿完成后,送至期刊出版社等待发表的文献,或是已投稿但未被审核接受的文章,也可以是未投稿至任何期刊的论文。后印本是指经过同行评审,并多次校对已经正式出版的文章。在传统的期刊出版过程中,预印本是研究成果正式发表前同行之间进行非正式交流的基本手段。预印本文献具有开放程度高、时效性强、学术性强、被引率高、费用低等特征。1991年8月,美国洛斯阿拉莫斯国家实验室(Los Alamos National Laboratory)的物理学家 Paul Ginsparg 建立的 arXiv.org 是第一个电子预印文本库。近年来,我国也出现了一些预印本文献,如中国科学技术信息研究所与国家科技图书文献中心共同开发的“中国预印本服务系统”(网址:<http://preprint.nstl.gov.cn/newprint/index.jsp>)、教育部科技发展中心主办的“中国科技论文在线”(网址:<http://www.paper.edu.cn>)等。

(13) 多媒体(Multimedia):是近年出现的一种新型的文献载体。它将声音、图像、文字、数据录入光盘,通过计算机实现重放或检索,因此具有前几种文献载体的优点,发

展特别迅速。

除了上述介绍的 13 种主要文献类型之外,还有新闻稿、统计资料等类的文献。在上 13 类文献中,一般把图书、期刊、报纸、电子预印本文献作为普通文献,其他均列入特种文献。

另外,按文献内容的学科属性划分分为哲学文献、社会科学文献、自然科学文献、综合科学文献及专科文献;按文献产生的时代阶段划分分为古代文献(1840 年前)、近代文献(1911 年前)、现代文献(1949 年 10 月前)、当代文献(1949 年 10 月以后)。

1.2 文献信息检索

如果说图书馆是一座知识的宝藏,那么“文献信息检索”就是打开这座宝藏的金钥匙。文献信息检索的发展经历了萌芽、发展和现代化三个阶段。

18 世纪中期,人们为了便于进行学术交流和文献信息的交流,不少学术团体先后创办了自己的学术刊物和检索性刊物,如德国 1769 年创刊的《各学院优秀外科论著摘要汇编》,是国外最早创办的一种检索性刊物。

到 20 世纪中期,随着科学技术的发展和需要,逐渐形成了较为完善的传统手工检索工具——目录、索引和文摘。

20 世纪的 60 年代后,形成了电子计算技术、光学缩微技术和网络通信技术三位一体的现代化文献信息检索手段,使文献信息检索发展到了一个新的阶段。

1.2.1 文献信息检索的涵义、类型和基本原理

1. 文献信息检索的涵义

文献信息是指关于文献的线索和文献中记录着的信息。文献信息检索是信息检索和文献检索两个概念的统一。

文献检索的概念有狭义和广义之分。狭义的文献检索是指依据一定的方法,从已经组织好的大量有关文献集合中,查找并获取特定的相关文献的过程。这里的文献集合,不是通常所指的文献本身,而是关于文献的信息或文献的线索。如果真正要获取文献中所记录的信息,那么还要依据检索所取得的文献线索索取原文。广义的文献检索包括存储和检索两个过程。存储是指工作人员将大量无序的文献信息集中起来,根据文献源的外表特征和内容特征,经过整理、分类、浓缩、标引等处理,使其系统化、有序化,并按一定的技术要求建成一个具有检索功能的工具或检索系统,供人们检索和利用。而检索是指运用编制好的检索工具或检索系统,查找出满足用户要求的特定文献。

信息检索是指依据一定的方法,从已经组织好的有关大量信息集合中,查出特定的相关信息的过程。

2. 文献信息检索的类型

根据检索(查找)对象的不同,文献信息检索分为文献检索、事实检索、数据检索和概念检索。

文献检索(Document Retrieval):文献检索所检索到的是关于文献线索的文献全文,它回答的是诸如“关于铁路大桥有哪些文献”之类的问题。

文献线索包括文献题目、著者、来源或出处、文摘等项目。文献线索检索指从一个文

献集合中找出专门文献的活动、方法与程序,是利用检索系统的工具查找文献线索,获取情报信息的过程,本质是文献需要与文献集合的匹配。例如,“关于自动控制系统有什么参考文献?”这就需要我们根据课题要求,按照一定的检索标识(如主题词、分类号等),从所收藏的文献中查出所需要的文献。

数据检索(Data Retrieval):是以数据为检索对象,从已收藏数据资料中查找出特定数据的过程。例如,查马拉雅山有多高,杭州六和塔建于何年等。

事实检索(Fact Retrieval):即通过对存贮的文献中已有的基本事实,或对数据进行处理(逻辑推理)后得出新的(即未直接存入或所藏文献中没有的)事实过程。例如:本学年成绩优秀多少;某同类产品中哪种牌号的销量最大。

概念检索(Concept Retrieval):就是查找特定概念的含义、作用、原理或使用范围等解释性内容或说明。

数据、事实和概念检索所得到的是能够确切解答问题的信息,或者说是文献中的具体信息。文献线索检索则是要检索出包含所需要信息的文献,其结果是与某一课题有关的若干篇论文、书刊的来源出处以及收藏地点等。

文献检索是最典型、最重要和最常用的文献信息检索。掌握了文献检索的方法就能以最快的速度,在最短的时内,以最少的精力了解前人和别人取得的经验和成果。

3. 文献信息检索系统的构成

无论采用什么手段对何种类型的检索系统进行文献信息检索,其检索系统必须具备四大要素:检索文档、检索设备、系统规则、作用于系统的人。

1) 检索文档

检索文档就是经过序列化处理并附有检索标识的信息集合。例如,手工检索系统使用的检索文档是由卡片式目录、文摘、索引所构成的系统;计算机检索系统使用的是存储在磁性或光性介质上的目录、文摘、索引或全文以及多媒体信息所构成的数据库。

2) 检索设备

检索设备即用以存储信息和检索标识,并实现信息检索标识与用户需求特征的比较、匹配和传递的技术手段,即检索所需的硬件环境。在手工检索系统中指印刷型检索工具,在计算机检索系统中包括各种类型的主机、终端、计算机外围设备和网络通信传输设备。

3) 系统规则

系统规则是用以规范信息总采集分析、标引著录、组织管理、检索与传输等过程的各项标准体系,例如检索语言,著录规则,检索系统构成与管理,信息传输与控制标准,输出标准等规则。

4) 作用于系统的人

作用于系统的人包括信息用户,信息采集分析、信息标引员,系统管理与维护员,检索服务人员等。

4. 文献信息检索的基本原理

文献信息检索的基本原理是:通过对大量的、分散无序的文献信息进行搜集、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文献信息源。其中存储是为了检索,而检索又必须先进行存储。文献信息的存储和检索的全过程可用图 1.1 表示。

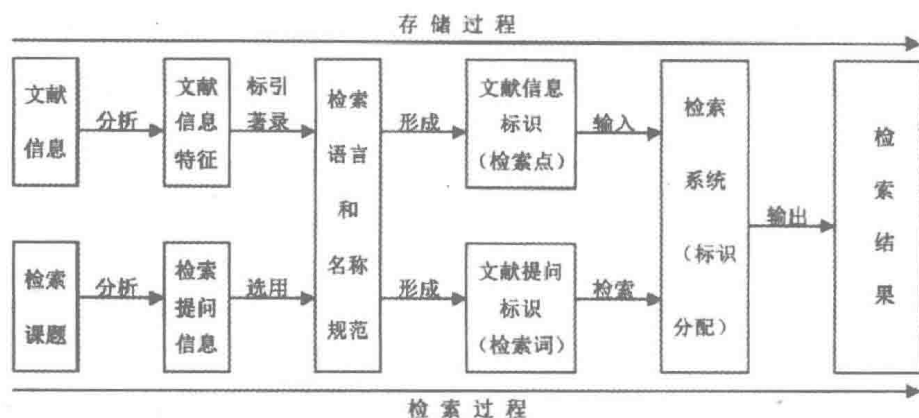


图 1.1 文献信息检索基本原理示意图

存储的过程,主要对信息源进行标引,将其外表和内容的特征(如文献的标题、作者、来源和主题等)用特定的检索语言转化为一定的标识(如主题词、分类号和类目名称等),再将这些标识按一定的顺序编排后输入检索系统,从而为检索提供有规可循的途径。

5. 文献信息检索与专家评审及科技查新的区别

文献信息检索是针对具体课题需要,仅提供文献线索或原文,对课题不进行分析和评价。专家评审主要是依据专家本人的专业知识、实践经验、对事物的综合分析能力以及所了解的专业信息,对被评对象的创造性、先进性、新颖性、实用性等做出评价。评审专家的作用是一般科技情报人员无法替代的,但具有一定程度的个人因素。

科技查新是文献检索和情报调研相结合的情报研究工作,它以文献为基础,以文献检索和情报调研为手段,以检出结果为依据,通过综合分析,对查新项目的新颖性进行情报学审查,写出有依据、有分析、有对比、有结论的查新报告。也就是说查新是以通过检出文献的客观事实来对项目的新颖性做出结论。因此,查新有较严格的年限、范围和程序规定,有查全、查准的严格要求,要求给出明确的结论,查新结论具有客观性和鉴证性,但不是全面的成果评审结论。这些都是单纯的文献信息检索所不具备的,也有别于专家评审。

1.2.2 检索语言及其类型

1. 检索语言的概念

把文献存储与检索联系起来,把标引(及著作)人员与检索人员沟通起来的思想工具就叫做检索语言。

文献检索语言是一种人工语言,用于各种检索工具的编制和使用,并为检索系统提供一种统一的、作为基准的、用于信息交流的一种符号化或语词化的专用语言。因其使用的场合不同,检索语言也有不同的叫法。例如在存储文献的过程中用来标引文献,叫标引语言;用来索引文献则叫索引语言;在检索文献过程中则为检索语言。

2. 检索语言的类型

检索语言的种类很多。按学科范围分有:综合性语言、专业性语言;按标识的使用方