



面向“十二五”高等教育课程改革项目研究成果

DIANZI WENXIAN

电子文献检索教程

JIANSUO JIAOCHENG

主编 胡光林 李雪萍

 北京理工大学出版社
BEIJING INSTITUTE OF TECHNOLOGY PRESS

面向“十二五”高等教育课程改革项目研究成果

电子文献检索教程

主 编 胡光林 李雪萍

副主编 胡 春 冯 凯 彭秋发



北京理工大学出版社

BEIJING INSTITUTE OF TECHNOLOGY PRESS

内容简介

本书系统地讲述了信息检索的相关知识和现代信息环境下信息检索的基本理论, 选择介绍了常用的中外工具书、电子文献资源数据库及重要网络信息资源, 并结合对这些检索工具的介绍, 采用了大量图解和实例讲解信息检索的方法和技巧, 以培养学生的信息意识和检索能力。本书主要内容有: 文献信息与信息意识、信息检索基本原理与技术、Internet 信息资源及其检索、常用中外文摘的使用、常用中外数据库、常用参考工具书、特种文献信息检索、古籍文献检索、信息分析与学术论文写作等。本书知识系统, 内容新颖, 教学目的明确, 适用性强, 可作为高等院校信息检索课程的教学用书, 也可作为广大读者提高信息检索能力的实用参考书。

版权专有 侵权必究

图书在版编目 (CIP) 数据

电子文献检索教程/胡光林, 李雪萍主编. —北京: 北京理工大学出版社, 2010. 8

ISBN 978 - 7 - 5640 - 3314 - 9

I. ①电… II. ①胡… ②李… III. ①互联网络 - 情报检索 - 高等学校: 技术学校 - 教材 IV. ① G354. 4

中国版本图书馆 CIP 数据核字 (2010) 第 123465 号

出版发行 / 北京理工大学出版社

社 址 / 北京市海淀区中关村南大街 5 号

邮 编 / 100081

电 话 / (010) 68914775 (总编室) 68944990 (批销中心) 68911084 (读者服务部)

网 址 / [http://www. bitpress. com. cn](http://www.bitpress.com.cn)

经 销 / 全国各地新华书店

印 刷 / 北京市通州富达印刷厂

开 本 / 710 毫米 × 1000 毫米 1/16

印 张 / 21. 25

字 数 / 397 千字

版 次 / 2010 年 8 月第 1 版 2010 年 8 月第 1 次印刷

印 数 / 1 ~ 4000 册

定 价 / 36. 00 元

责任校对 / 陈玉梅

责任印制 / 边心超

图书出现印装质量问题, 本社负责调换

前 言

21 世纪是信息化、数字化、全球化的知识经济时代,计算机的产生和发展为科学技术插上了强有力的“翅膀”,互联网是其鲜明的特色标志和重要组成部分。随着网络技术的不断发展和电子文献资源的日益增多,海量规模的文献资源给广大读者带来了新的困惑——从“找不到资源”转变为“不知如何最方便、快捷地找到所需资源”。人们曾用“知识爆炸”来形容当今世界科学技术知识的增长,如何从这浩如烟海的信息中找出所需信息,就成为信息检索的重任。

伴随信息科技步入多媒体发展时期,计算机信息检索必然地发展到了应用与发展阶段,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,也是时代的需要。

信息检索是一门关于信息资源的存储、整理和查找的理论及方法的学问。本书是介绍信息检索的原理和方法及中外最常用的检索工具的教材。结合时代的特点,本书重点介绍电子资源文献检索方法及技巧,考虑到许多大学都有文史专业,还设置了古籍资料的检索章节,以适应教学的需要。

要学好信息检索这门课程,需要了解本课程的性质和任务、教学目的、课程意义和作用、教学方法和要求等。

课程性质和任务:信息检索课程是旨在培养大学生的信息意识、掌握从文献中获取信息知识的一门科学方法课程。本课程的任务是使学生了解各自专业及相关专业信息检索的基本知识,学会常用检索工具与参考工具的使用方法,懂得如何获得与利用信息资源,增强自学能力和研究能力。

教学目的:通过教学使学生掌握信息检索的基本原理和基本技能,提高学生的信息素质和综合创新能力,增强利用信息进行分析问题和解决问题的能力。

课程意义和作用:有利于减少科研活动的重复劳动,提高科研成功率;有利于节省科研时间与科研开支,提高科研效益;有利于为社会提供服务,提高社会效益;有利于开发信息资源,提高经济效益。

教学方法和要求:以多媒体课堂教学为主,以上机实习为辅的教学模式。要求学生深刻理解授课内容,不仅要重视理论知识学习,更要重视实践操作;要求学生理解和掌握信息检索的知识,举一反三,适应知识经济时代信息化的新要求。

目前,许多大学都开设了信息检索及其相关课程,部分学校还专门为研究生开设了这门课程,这足以看出大学教育对信息检索课程的重视。本教材主要为高等院校本科生的教学而编写,学时为30~40。在教学过程中,可以根据专业性质的不同在内容上进行适当的删减和取舍。

本书由长期从事信息检索教学的教师结合多年教学实践经验编写而成,其中第一章由冯凯、喻逸君编写;第二章和第三章由冯凯、郑颖佳编写;第四章由胡光林、高宇编写;第五章由王嵩、危志明编写;第六章由胡春、高宇编写;第七章由张文华、杨福来编写;第八章由胡春、张兰编写;第九章由李雪萍编写。全书由胡光林、李雪萍、胡春、冯凯、彭秋发等统一组织、策划和统稿。

由于编者水平有限,书中难免有疏漏和不足之处,恳请读者批评指正。

编 者

目 录

第一章 文献信息与信息意识	(001)
第一节 信息、知识与文献	(001)
第二节 文献的特点与特征	(006)
第三节 文献的资源类型	(011)
第四节 信息意识	(018)
思考题	(023)
第二章 信息检索基本方法和基本技术	(024)
第一节 信息检索原理、类型与方式	(024)
第二节 信息检索语言与工具	(027)
第三节 信息检索的方法、途径与程序	(033)
第四节 计算机信息检索	(042)
思考题	(051)
第三章 Internet 信息资源及其检索	(053)
第一节 Internet 信息资源检索	(053)
第二节 搜索引擎及其使用	(058)
第三节 常用中文搜索引擎	(064)
第四节 常用外文搜索引擎	(082)
思考题	(089)
第四章 常用中外文摘的使用	(090)
第一节 《全国报刊索引》	(090)
第二节 美国《科学引文索引》	(096)
第三节 美国《工程索引》	(106)

思考题	(112)
第五章 常用中外数据库	(113)
第一节 中国知网 CNKI	(113)
第二节 重庆维普数据库	(122)
第三节 人大复印资料数据库	(130)
第四节 万方数据资源系统	(133)
第五节 超星电子图书数据库	(139)
第六节 SpringerLink 数据库	(142)
第七节 EBSCO 全文数据库	(147)
第八节 其他外文数据库	(153)
思考题	(157)
第六章 常用参考工具书	(159)
第一节 数据与事实检索概述	(159)
第二节 参考工具书概述	(159)
第三节 常用参考工具书简介	(162)
第四节 网络工具书	(185)
思考题	(194)
第七章 特种文献及其检索	(195)
第一节 会议文献检索	(195)
第二节 学位论文检索	(207)
第三节 标准文献检索	(221)
第四节 专利文献检索	(234)
思考题	(253)
第八章 古籍资料的检索	(254)
第一节 书目的利用	(254)
第二节 类书的利用	(259)
第三节 政书的利用	(263)
第四节 索引的利用	(268)
第五节 数字化资源的利用	(280)
思考题	(288)

第九章 信息分析与学术论文写作	(290)
第一节 信息分析概述	(290)
第二节 信息分析基本程序	(294)
第三节 信息分析研究方法及评价	(296)
第四节 学术论文的撰写	(303)
思考题	(318)
附录 中国图书馆图书分类法 (简表)	(319)
参考文献	(327)

第一章

文献信息与信息意识

第一节 信息、知识与文献

一、信息

1. 信息的概念

随着科学技术的飞速发展，现代社会已进入信息时代。通信技术、计算机技术和音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。信息的概念十分广泛，围绕信息而出现的信息资源、信息技术、信息系统、信息产业、信息化社会和社会信息化等相关术语不胜枚举。可见，信息的观点、概念和方法已经被政治、经济、科技、文化、生产等各个领域所接受和应用。

那么，信息（Information）究竟是什么呢？“信息”这个词最早见于我国唐代诗人李中《暮春怀故人》中“梦断美人沉信息，目穿长路倚楼台”的诗句，这里的信息就是消息的意思，中国古朴的“信息”的含义通常理解为信号和消息，随着科技的不断发展和人类的不断进步，信息的概念也不断演变深化。

至今为止，信息还没有一个统一的定义，不同的人站在不同的角度就有不同的理解和解释。例如，“信息论”的创立者申农（C. E. Shannon）着眼于对信息作用的描述，把信息定义为“用于消除不定性（Uncertainty）的东西”。“控制论”的创立者维纳（N. Wiener）着眼于信息应用，他对信息的定义是：“信息是人们在适应外部世界并且反作用于外部世界的过程中同外部世界进行交换的内容的名称。”《辞海》中将信息解释为：“一是指音讯；消息。二是指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，通常须通过处理和分析来提取。”我国国家标准《情报与文献工作词汇基本术语》中记载：“信息是物质存在的一种方式、形态和运动状态，也是事物的一种普遍属性，一般指数据、消息中所包含的意义可以使消息中所描述事件的不定性减少。”

综合各家定义的合理内核,本书把信息定义为: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变化中所具有特征性的内容的总称。不同事物具有不同的存在方式和运动规律,从而构成了各种事物的不同的特征,发出了不同的信息。

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。通俗地讲,信息就是通过信号带来的消息,信息普遍存在于自然界、社会和思维中,信息与物质、能量构成了现代社会的三大资源,成为了社会进步的强劲推动力。

现在,人类已走进信息社会的时代,信息作为一种战略性资源,已经成为发展科技、生产、经济、文化、教育的重要支柱之一。

了解了信息的概念和含义,接下来要了解什么是信息检索(Information Retrieval)。“信息检索”原称情报检索,萌芽于图书馆的参考咨询工作,20世纪50年代才固定成专用术语。随着信息爆炸和以计算机技术为核心的信息技术的迅速发展,信息检索的概念与类型都在发生新的变化。

有关信息检索的定义,目前国内外也有不同的表述,本书将其理解为:信息检索是指将信息按一定的方式组织和存储起来,并根据信息用户的需要找出有关的信息的过程,所以它的全称又叫信息的存储与检索(Information Storage and Retrieval),这是广义的信息检索。狭义的信息检索则仅指该过程的后半部分,即从信息集合中找出所需要的信息的过程,相当于人们通常所说的信息查寻(Information Search)。

而计算机信息检索,是指以计算机技术为手段,通过光盘和联机等现代检索方式进行信息检索的方法。

2. 信息的特点

信息按其性质分为自然信息和社会信息;按其载体可分为空气信息、电子信息、光纤信息和文献信息等。不论哪种类型的信息,都有着以下共同的特点。

1) 普遍性

信息既不是物质,也不是能量,它是客观事物普遍性的表征,信息是无处不在、无时不有的普遍现象。只要有事物存在,只要有事物的运动,就会有其运动的状态和方式,就存在着信息,因此,信息是普遍存在的。

2) 传递性

信息可以通过多种渠道、采用多种方式进行传递,通常把信息从时间或空间上的某一点向其他点移动的过程称为信息传递,信息传递要借助于一定的物质载体,实现信息传递功能的载体又称为信息媒介。没有物质作为载体,信息也就不能传递和交换。

3) 共享性

在信息的扩散或传播过程中,同一内容的信息在同一时间、同一地域可以被多个用户分享,其分享的信息量不会因分享用户的多少而受影响,原有的信息量也不会因这个过程而损失或减少,也不会因分享的次数而影响自身的信息源。

4) 扩散性

物质是普遍联系和相互作用的,作为表征物质相互联系和相互作用的信息,一方面同时或连续向其他物质系统扩散传播,另一方面也接受其他物质系统扩散的信息,而且,这样的相互扩散是无穷无尽的。

5) 时效性

现代社会中,信息的使用周期迅速缩短,信息的价值实现取决于及时地把握和运用信息。信息是活跃的、不断变化的。及时地获取有效的信息将获得信息的最佳价值,如时效性很强的天气预报、经济信息、交易信息、科学信息等。

6) 价值性

信息通过人脑思维或人工技术设备的综合、加工和处理,不断积累丰富,提高其质量和利用价值。信息的质量和价值,实际上是对客观事物属性反映的深度和真实程度的认识。虽然信息是人类的一种重要资源,但(其)只有被利用才会产生价值,否则,其价值或随时间的流逝而减少,或成为“信息垃圾”。

7) 可储性

信息可以收集、加工、整理、筛选、归纳、综合,形成如声音、文字、图像及用各种编码形式记录下来的数据,并可以通过记忆和各种载体来载荷,这就是信息的可储性。

二、知识

《辞海》中将“知识”(Knowledge)定义为“人们在社会实践中积累起来的经验”,并指出“从本质上说,知识属于认识的范畴”。国外有些学者认为知识是一种能够改变某些人或某些事物的信息。由此可见,知识是建立在信息的基础之上的,是人类通过信息对大自然以及人类本身进行挖掘、发现、分析、综合而创造出来的新的信息,是通过实践活动和大脑的思维而总结出的新的认识,是改造自然和人类本身必须有的信息活动。由此可以看出,知识就是信息的一部分,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与掌握,是人脑通过思维重新组合的、精简化的信息集合。人类知识存在的形式基本上有两类:一类是存在于人脑记忆中的知识,属于主观知识;另一类是用文字、图形、符号、视频、声频等技术手段记录在一定载体上的知识,属于客观知识,通常人们把这类知识称作文献信息。只有在人们的使用过程中,知识才体现出其价值,才成为具有实践意义的、真正的知识。

三、情报

情报为外来语,来自日语“情报”(じょうほう),主要指“信息、资讯、消息”。“情报”一词最早产生于军事领域,是战时关于敌情的报告,其在不同的历史时期有着不同的含义。但是,无论情报的内容与形式如何变化,其共同之处就是情报是指为一定目的搜集和传递的有特定效用的知识。它具有知识性、传递性和效用性3个基本属性。

情报的知识性是指情报的本质是知识,即任何学科领域的、以任何形式出现的情报都具有一定的知识内容。人们通过阅读图书、报刊,通过听广播,看电视、电影,通过参观、考察、调研等活动,随时都在接受、传递和利用大量的感性和理性的知识。这些知识就包含着人们所需要的情报。因此,情报是引起人们原有知识结构发生改变的知识,是活化了的的知识。

情报的传递性是指知识只有经过传递才能成为情报。正如上面所述,情报是活化了的的知识,这种使知识由静态进入动态而转变成活化知识的过程就包含着知识的传递,这种传递是通过语言文字、电磁波等各种情报载体来实现的。广播电视,计算机信息处理,激光、光缆、通信卫星,以及信息高速公路等先进技术的出现与发展,极大地扩展了情报的传递范围,加快了情报的传递速度。所以,情报就是流动的知识。

情报的效用性是指并非流动着的知识都是情报,只有那些能满足特定需要地流动着的知识才可以称之为情报。例如,每天通过广播传递的大量信息,是典型的流动的知识,但对于大多数人来说,这些广播内容只是消息,只有少数人利用广播的内容增加了知识或解决了问题。这部分被发挥了效用地流动着的知识才是情报。因此,情报是对特定的人而言的,对需要它的人而言是情报,对不需要的人而言,它只是信息与知识。所以,情报是一种动态的有用的知识。

四、文献

“文献”(Literature, Document)一词最早出现在《论语·八佾》中——子曰:“夏礼吾能言之,杞不足征也;殷礼吾能言之,宋不足征也。文献不足故也。足,则吾能征之矣。”大意是:夏朝礼制我能讲,但杞国不足以证明;殷朝的礼制我也能讲,但宋国不足以证明。因为这两国文献不足,否则我也可以证明,这段话表明了孔子论事有据,注重文献的精神。这里孔子只是说明了文献的足与不足可以作为能否讲解杞宋两国礼制的一个重要依据,而并未说明文献指什么。(宋)朱熹注释“文,典籍也,献,贤也”说明文献一词在古代的原意是指典籍和宿贤。所以当人们要了解往事,研究历史,讨论问题,不能单靠记载下来的

“文”。还必须请教那些阅历丰富、熟悉掌故的贤人，因此，“文献”一词自然也就包含了两方面的意思。到了现代，“文献”一词的概念有了很大的改变，其重要一点是将那些阅历丰富、熟悉掌故的贤人从文献概念中排除，即文献的含义侧重于“文”，“贤”的意义消失了。

《辞海》将文献定义为“原指典籍和宿贤，现指具有历史价值的图书文物资料，如历史文献。亦指与某一学科有关的重要图书资料，如医学文献”。根据中国国家标准《文献著录总则》（GB 3792.1—1983），其对文献的定义为：“文献是指记录有知识的一切载体。”《文献情报术语国际标准（草案）》（ISO/DIS5127）指出：“文献是指在存储、检索、利用或者传递记录信息的过程中，可作为一种单元处理的，在载体内、载体上或者依附载体而存储有信息或数据的载体。”综上所述，这点将文献定义为：文献就是用文字、图形、符号或声像、视频等技术手段记录知识信息的物质载体，或者说是固化在物质载体上的知识信息。

文献的构成主要有3个要素：第一，有一定的知识内容；第二，有用以保存和记录知识的技术手段；第三，有记录知识的物质载体。这3个要素是文献的内涵。不同的知识内容、不同的记录方式和不同的物质载体，可形成不同类型的文献，这就是文献概念的外延。

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，还包括古代的甲骨文、竹简、帛书等，以及当今的声像出版物、电子出版物和互联网上的信息。

五、信息、知识、情报及文献之间的关系

信息、情报、知识、文献4个概念既有区别又有联系，它们之间的逻辑关系是：知识、情报和文献的本质都是信息，它们都属于信息的一部分，但它们的范围大小又有所不同。知识属于信息这一大范畴，是信息的一部分，信息中的一部分经过人的大脑思维之后重新组织和整合，变成对人类有用的、精简化的信息集合，就成为了知识；情报是知识的一部分，知识中的一部分被传递着用于特定的目的时就变成了情报；文献是记录有知识的一切载体，是知识中的一部分被人类用特定的载体记载了下来而形成的，知识分为主观知识和客观知识，文献记载的是客观知识。这几个基本概念之间的关系如图1-1所示。

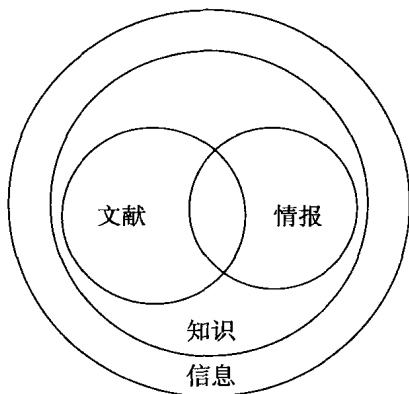


图 1-1 信息、知识、文献及情报之间的关系示意图

第二节 文献的特点与特征

一、文献的特点

随着科学技术的不断发展，当今社会已经进入信息时代，文献相应地具有以下几个特点。

1. 数量急剧增长

随着科学技术和社会生活的飞速发展，人类知识总量猛增。文献作为存储、传播知识的载体文献，其数量也随着知识量的增加而激增，而且增长速度很快。以我国为例，1949—1954 年，我国出版图书 21 809 种。2006 年我国出版图书 233 971 种，其中新版图书 130 264 种；出版期刊 9 468 种，平均期印数 16 435 万册，总印数 28.52 亿册；出版报纸 1 938 种，平均期印数 19 703.35 万份，总印数 424.52 亿份；音像制品及电子出版物为 40 913 种。共 6.4 亿张；现存古籍约 10 万多种。目前，全世界每年出版图书 100 多万种，期刊 50 多万种，其他文献信息资料 500 多万种；发表科学论文大约 500 万篇，平均每天 14 000 篇左右，每 35 秒就有 1 篇论文发表，不到 1 分钟就有 1 本新书问世，每小时出现 20 多项技术发明，每天约有 40 亿个信息单位的信息量向全世界发送。表 1-1 所示为 1990—2005 年我国图书、期刊、报纸出版总量统计。

表 1-1 1990—2005 年我国图书、期刊、报纸出版总量统计

年 份	总印张/亿印张	折合用纸量/万吨
1990	463.0	108.6
1991	526.32	123.5
1992	581.89	136.23
1993	633.61	147.95
1994	671.77	156.67
1995	743.42	173.21
1996	820.92	191.31
1997	897.11	208.81
1998	993.48	231.02
1999	1124.81	261.33
2000	1276.08	296.05
2001	1445.96	335.27
2002	1630.21	377.95
2003	1806.93	418.59
2004	2100.9	486.19
2005	2231.67	523.45

文献数量的激增，一方面表明文献信息资源的丰富，但另一方面也产生了“文献信息污染”，给查找相关文献带来了困难。例如，一名化学家每周阅读 40 小时，浏览世界上一年内发表的有关化学方面的论文和著作就要 48 年。因此，需要人们选择、利用所需的文献，获取所需信息，学会检索和利用相关文献资源。

2. 分布既集中又分散

随着现代科学技术的日益综合与细化，各学科之间的严格界限日趋淡化，学科之间的相互联系、交叉渗透逐渐增强。这使文献的分布呈现出既集中又分散的现象。1931 年，著名文献学家布拉德福 (B. C. Bradford) 首先揭示了文献集中与分散规律，发现某时期某学科 $1/3$ 的论文刊登在 3.2% 的期刊上；1967 年联合国教科文组织研究了二次文献在期刊上的分布，发现 75% 的文献出现在 10% 的期刊中；1971 年，SCI 的创始人加菲尔德 (E. Garfield) 统计了参考文献在期刊上的分布情况，发现 24% 的引文出现在 1.25% 的期刊上；等等。这些研究都表明期刊存在“核心效应”，从而衍生了“核心期刊” (Core Periodicals) 的概念。

核心期刊是指那些刊载论文学术质量高、科技信息含量大，刊物学术影响广

泛,被公认为能反映某种学科(专业)科学研究的前沿动态、学术水平和发展方向的学术期刊。

例如,利用英国《科学文摘》统计了“日本光纤通信”方面的论文322篇,共发表在66种杂志上。发现有107篇文章发表在3种杂志上(核心期刊),约有33%的文章发表在另12种期刊上(相关期刊),还有近1/3的文章发表在另51种期刊上(边缘期刊)。文献的这种分布现象是普遍存在的,它提示人们在收集文献信息时,应首先选择本专业的核心期刊,但也不能忽视其他相关期刊。

我们可以使用有关的工具来了解某专业的核心期刊,例如由北京大学图书馆等单位合编的《中文核心期刊要目总览》(2008年版)、《国外科学技术核心期刊总览》(2004年版)、《国外人文社会科学核心期刊总览》(2008年版)等。

3. 失效加快

目前,知识更新周期不断缩短,文献“老化”不断加剧。据悉,各类文献的平均寿命为:图书10~20年,科技报告10年,学位论文5~7年,期刊论文3~5年,标准文献5年,产品样本3~5年。据欧美一些发达国家统计,文献的平均寿命5~7年。这就要求图书界必须加快文献信息的传递速度,才能适应生产技术和科学研究发展的需要。同时,各种类型的文献有不同的有效使用时间,因此,各学科的文献寿命不同,各国科技发展水平不同,相应的文献寿命也不相同。

4. 内容交叉重复

现代科学技术综合交叉、彼此渗透的特点,导致知识的产生和文献的内容也相互交叉、彼此重复。具体表现如下。

1) 同一内容的文献以不同文字发表

据统计,当前世界上每年翻译图书约占图书出版总量的10%;一些重要的核心期刊被译成多种文字在不同国家出版;同一项发明可以向多个国家申请专利,使专利说明书的内容重复现象严重,据世界知识产权组织统计,世界各国每年公布的专利说明书的重复率达65%~70%。

2) 同一内容的文献以不同形式出版

据调查,会议论文有40%会后经过整理、修改发表在期刊上;有相当数量的科技报告后来发表在期刊上,如美国的NASA报告有80%、AD报告有60%、美国科学基金会报告有95%、美国农业部的科技报告有80%,既出版单行本又发表在期刊上。随着新型载体文献的普及应用,许多文献既出版印刷型的,又有电子版等。如《中国大百科全书》、《不列颠百科全书》、《工程索引》、《科学文摘》、《科学引文索引》等,既有印刷版,又有电子版。

3) 许多畅销书内容雷同, 选题重复

在激烈的商业竞争中, 再版、改版文献数量不断增多, 如 2006 年我国出版图书 233 971 种, 其中新版图书 130 264 种, 新版图书仅占图书出版总量的 55.6%。

5. 载体及语种增多

随着声、光、电、磁等技术和新材料的广泛应用, 新型文献载体不断涌现。传统的纸张型文献已失去了一统天下的局面, 多种载体的文献相互依存、相互补充、共同发展已成为发展趋势。新型的非纸张型文献或增大了信息存储密度, 延长其保存时间; 或加快了信息传递与检索速度, 实现了资源共享; 或令人闻其声、见其形, 使文献家族日臻繁荣。计算机技术、电子技术、远程通信技术、光盘技术、视听技术、网络技术等, 构成了信息的现代传播技术。联机检索、交互式图文检索、电子原文传递等现代化信息传播方式已进入实用阶段。信息检索已发展到网络化阶段, 人们可以利用分布全球的互联网, 多途径、多选择、多层次地检索所需文献。

各国文献所用语种不断增多。据 ISDS (International Serials Data System) 1991 年报道, 世界上连续出版物使用的语种多达 144 种, 常用的就有 10 多种, 其中英文约占 42%。

二、文献的特征

文献检索是通过将表示检索要求特征的信息与存储 (记录) 在文献数据库中的文献特征信息作相符性比较来实现的。检索者的检索要求不外乎有两种: 一是要查出具具有已知文献外表特征的文献, 如由书名、作者途径查找相关文献; 二是要查出具有所需内容特征的文献, 如根据某学科要求编写的一本教材或围绕某一课题收集有关的资料。为了满足检索者这两种检索要求, 信息工作者要用特定语言表达分析出的文献的属性或特征, 这就是标引, 标引人员在编制检索工具 (或系统) 时, 按照文献的外表特征和内容特征分别进行标引, 以形成满足检索者这两种需求的多种不同的检索途径。

1. 外表特征

文献的外表特征通常包括题目、作者、作者工作单位、专利和科技报告, 以及专利号或报告号等, 这些可以表征一篇特定文献的特征, 它可以在文献的封面或扉页, 即不打开书本, 或不看文献的具体内容就可以确定一篇文献的相关信息。根据文献的外表特征, 其查准率 (Precision Ration) 很高, 查准率能反映每次从该系统文献库中实际检出的全部文献中有多少是相关的。

也就是说, 查准率是指检出的相关文献量与检出文献总量的比率, 是衡量信