

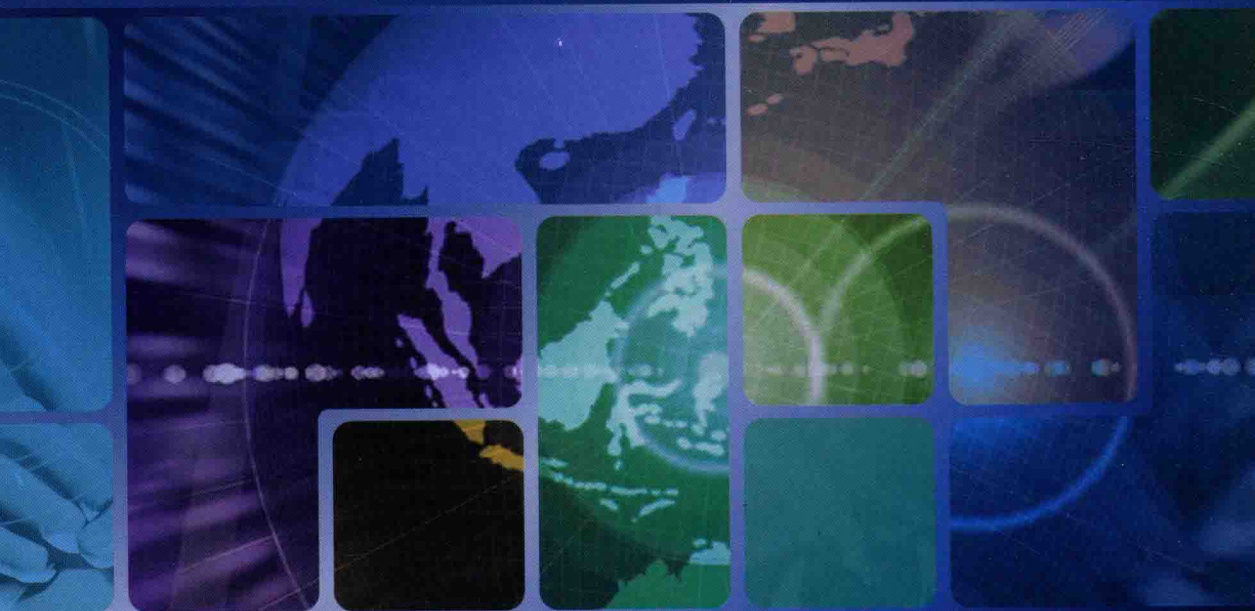


普通高等教育“十二五”规划教材

信息检索

(第二版)

杜 伟 主编



科学出版社

普通高等教育“十二五”规划教材

分 册 内 容

信息检索

(第二版)

杜 伟 主编

普通高等教育“十二五”规划教材

《信息检索》是“十二五”规划教材，由杜伟主编。

本书可作为高等院校图书馆学、情报学、档案学、信息管理等专业

及相关专业的教材，也可供从事信息工作的有关人员参考。

本书共分五章，第一章为绪论，第二章为信息检索基础，第三章为

信息检索方法，第四章为信息检索工具，第五章为信息检索评价。

本书可作为高等院校图书馆学、情报学、档案学、信息管理等专业

及相关专业的教材，也可供从事信息工作的有关人员参考。

科学出版社

北 京

内 容 简 介

本书立足于介绍现代信息检索的方法与技术,在对信息检索基础知识、信息检索工具基础知识进行全面阐述的基础上,重点介绍国内外重要电子信息资源、网络信息资源的检索与利用,也对部分重要印刷型资源的检索进行了简单介绍,为读者提供了较全面的信息检索知识。

本书是一本工具型书籍,提供的信息检索学科知识范围广,内容尽量以最新版本为主,是一本适用性强的学习信息检索方法和技巧的图书。

本书既可以作为师范院校、综合性院校本科生、研究生的教材,又可以作为教学、科研及各界信息用户的参考工具书。

图书在版编目(CIP)数据

信息检索/杜伟主编. —2版. —北京:科学出版社,2015.3
普通高等教育“十二五”规划教材
ISBN 978-7-03-047444-5

I. 信… II. 杜… III. 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(2016)第 043596 号

责任编辑:毛 莹 / 责任校对:桂伟利
责任印制:霍 兵 / 封面设计:迷底书装

科学出版社 出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

新科印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2009 年 2 月第 一 版 开本:787×1092 1/16

2016 年 3 月第 二 版 印张:14 1/2

2016 年 3 月第 11 次印刷 字数:371 000

定价:32.00 元

(如有印装质量问题,我社负责调换)

《信息检索》编委会

主 编 杜 伟

副主编 王世慧 王爱国

主 审 夏亚云

编 者 杜 伟 王 君 王世慧

王爱国 赵炜霞 胡健飞 夏亚云

序

依据 1984 年教育部下达的[84]高教字 004 号文件《关于在高等学校开设文献检索与利用课的意见》,许多高校开设了文献检索或信息检索之类的必修课、公共课或选修课。经过 30 余年的检索课教学实践与探索,信息检索课程正逐步发展成为一门具有一定理论深度和知识广度的课程。相当多的院校已基本上形成了新生入学时的“入馆”教育、本科生的文献检索课程教育、研究生的专业信息资源查找和利用教育的教学体系。进入 21 世纪以来,一些高校适时地开设了网络信息检索与利用的课程。

在信息时代,社会经济全面信息化,劳动者是拥有丰富的知识和信息、掌握着高技术的现代人,劳动对象主要是信息和知识,劳动工具是由计算机控制的高度信息化和智能化的机器和现代工具,建立在知识和信息的生产、分配和使用基础之上的知识经济和信息经济将成为社会经济的主体,信息产业和知识产业将成为主导和支柱产业。信息环境的变迁、信息用户需求的变化以及信息资源的增多,使良好的信息素质成为现代人的必备素质。1989 年,美国图书馆协会就对信息素质做了定义。总之,在信息时代里具有信息素质的人应具备 6 个方面的信息技能:能知道何时需要信息,能鉴别解决疑难问题所需的信息,能查找出所需要的信息,能评价信息,能组织信息,能有效地利用信息解决遇到的问题。

完整的信息素质应包括信息意识、信息能力和信息道德。信息意识是指对知识信息重要性的认识和对知识信息的敏感性;信息能力也就是上面所说的信息技能,其重点和基础是信息检索能力;信息道德简要地说,是指人们在信息活动中应遵循的道德规范。高校信息检索课程的建设和发展,对培养学生这三方面的意识有着积极的作用,特别是培养学生的信息能力。

20 世纪 90 年代以来,随着通信技术、网络技术和与计算机相结合的现代技术的不断发展,新的知识内容、检索手段不断产生,传统的媒体和检索工具、检索方式在不断发生变化,人们的信息意识、信息观念也在不断更新,这为海量信息的开发和利用提供了便利条件,也为信息的发布和分享提供了外部环境,然而也导致信息的查找和使用上的困难。数字化、网络化信息的分散、无序和动态变化,以及信息的庞杂同用户特定需求之间的矛盾也给人们检索和利用带来障碍和不便。鉴于此类问题的存在及信息检索的发展需求,作者特地编写了《信息检索》一书,并不断更新内容和版本。

该书由 13 章构成,立足于介绍现代信息检索的方法与技术,在对信息检索基础知识、信息检索工具基础知识进行全面阐述的基础上,重点介绍国内外重要电子信息资源、网络信息资源的检索与利用,也对部分重要印刷型资源的检索进行了简单介绍,为读者提供了较全面的信息检索知识。

该书观点新颖,资料翔实,系统性、严谨性和科学性是该书保证高质量的前提,这也充分说明了作者在信息检索领域具有深厚而扎实的学术功底。他们的责任意识、大局意识和精品意识是应该加以肯定的。

该书作者长期以来一直从事信息开发与服务、科技信息查新及研究、信息检索和教学等工作,并出版过专著,同时他们都有高级技术职称,因此由他们编写的这本教科书是值得信赖的。该书的问世,对信息资源的开发利用,对大学生的信息素质教育和创新能力的培养将发挥积极作用。当然,由于时间紧、任务重,书中的缺点与不足之处在所难免,尚需从事信息检索教学的同仁匡正。

本书编委会主任 王 健 教授 中国矿业大学北京图书馆馆长 王 健 教授 2016 年 1 月于济南

本书编委会成员 王 健 教授 中国矿业大学北京图书馆馆长 王 健 教授 2016 年 1 月于济南

本书编委会成员 王 健 教授 中国矿业大学北京图书馆馆长 王 健 教授 2016 年 1 月于济南

本书编委会成员 王 健 教授 中国矿业大学北京图书馆馆长 王 健 教授 2016 年 1 月于济南

本书编委会成员 王 健 教授 中国矿业大学北京图书馆馆长 王 健 教授 2016 年 1 月于济南

本书编委会成员 王 健 教授 中国矿业大学北京图书馆馆长 王 健 教授 2016 年 1 月于济南

本书编委会成员 王 健 教授 中国矿业大学北京图书馆馆长 王 健 教授 2016 年 1 月于济南

本书编委会成员 王 健 教授 中国矿业大学北京图书馆馆长 王 健 教授 2016 年 1 月于济南

本书编委会成员 王 健 教授 中国矿业大学北京图书馆馆长 王 健 教授 2016 年 1 月于济南

本书编委会成员 王 健 教授 中国矿业大学北京图书馆馆长 王 健 教授 2016 年 1 月于济南

本书编委会成员 王 健 教授 中国矿业大学北京图书馆馆长 王 健 教授 2016 年 1 月于济南

本书编委会成员 王 健 教授 中国矿业大学北京图书馆馆长 王 健 教授 2016 年 1 月于济南

本书编委会成员 王 健 教授 中国矿业大学北京图书馆馆长 王 健 教授 2016 年 1 月于济南

本书编委会成员 王 健 教授 中国矿业大学北京图书馆馆长 王 健 教授 2016 年 1 月于济南

本书编委会成员 王 健 教授 中国矿业大学北京图书馆馆长 王 健 教授 2016 年 1 月于济南

本书编委会成员 王 健 教授 中国矿业大学北京图书馆馆长 王 健 教授 2016 年 1 月于济南

本书编委会成员 王 健 教授 中国矿业大学北京图书馆馆长 王 健 教授 2016 年 1 月于济南

本书编委会成员 王 健 教授 中国矿业大学北京图书馆馆长 王 健 教授 2016 年 1 月于济南

前 言

随着科学技术的不断进步和新信息技术的层出不穷,人们的信息意识、信息观念以及获得信息的手段也在不断地更新变化。

现在高等教育的职能已经不再是单纯的知识传授,而是把学生培养成信息素养人。信息检索的意识和能力是考察一个人信息素养的重要指标。培养学生的信息检索意识和能力对于提高学生的自学能力和独立研究能力有很大的益处。

随着网络的不断发展,网络信息资源越来越丰富。我们经常遇到这样的问题:怎样在网上找到相关信息?有哪些工具可以利用?这些网络工具或者电子资源如何使用?等等。所以,本书立足于介绍现代信息检索的方法与技术,在对信息检索基础知识、信息检索工具基础知识进行全面阐述的基础上,重点介绍国内外重要电子信息资源、网络信息资源的检索与利用,也对部分重要印刷型资源的检索进行了简单介绍,为读者提供了较全面的信息检索知识。

随着网络技术的发展,光盘数据库已不适应社会发展的需要,本书对部分内容作了调整,删除了光盘数据库检索,对网络数据库检索介绍得更加详细,突出实用性。

本书由济南大学图书馆信息素质教育教研室编写,杜伟负责全书的统稿工作,夏亚云研究馆员对全书进行了认真的审校,济南大学副校长蔡先金教授为本书作序。

由于水平所限,难免有疏忽或不足之处,望读者和同行指正。

编 者

2016年1月

目 录

序

前言

第1章 概论	1	5.4 印刷版文摘正文编排及著录格式	59
1.1 信息检索的意义	1	5.5 辅助索引	61
1.2 信息资源的类型	3	5.6 检索词表	62
1.3 文献信息的类型	4	5.6.1 工程主题词表	62
第2章 信息检索基础知识	10	5.6.2 工程索引叙词表	65
2.1 检索语言	10	5.7 来源出版物目录	67
2.1.1 检索语言基础知识	10	5.8 附表	68
2.1.2 检索语言的作用	11	5.9 网络检索系统	68
2.1.3 检索语言的类型	11	第6章 美国《化学文摘》	74
2.2 信息检索的功能和技术	18	6.1 概况	74
2.2.1 信息检索功能	18	6.2 网络版数据库——SciFinder	75
2.2.2 信息检索技术	20	第7章 美国《科学引文索引》	83
2.3 信息检索的途径和方法	27	7.1 概况	83
2.3.1 信息检索的途径	27	7.2 出版物类型	83
2.3.2 信息检索的方法	29	7.3 网络数据库检索——WOS系统中的SCI检索	84
第3章 检索工具基础知识	32	第8章 美国《社会科学引文索引》和《艺术与人文科学引文索引》	90
3.1 检索工具	32	8.1 概况	90
3.2 缩写	35	8.2 网络数据库检索——WOS系统中的SSCI和A&HCI检索	91
3.3 音译	37	第9章 专利信息检索	92
第4章 英国《科学文摘》	40	9.1 知识产权	92
4.1 概况	40	9.2 专利和专利文献	92
4.2 结构	42	9.3 专利信息检索的目的、类型和途径	98
4.3 主题词表	42	9.4 国际专利分类表	100
4.4 文摘著录格式	44	9.5 国际外观设计分类表	102
4.5 索引体系	48	9.6 中国专利文献	103
4.6 Web of Science INSPEC数据库检索	52	9.7 国内专利信息检索系统选介	107
第5章 美国《工程索引》	56		
5.1 概况	56		
5.2 出版形式	56		
5.3 报道范围	58		

9.7.1 国家知识产权局专利检索系统	107	12.5 CASHL 数据库	148
9.7.2 中国专利信息网	110	12.6 国家科技图书文献中心 数据库	150
9.7.3 中国知识产权网检索系统	110	12.7 万方数据资源系统	155
9.8 国外专利信息检索系统选介	111	12.8 读秀知识库	161
9.8.1 WIPO 国际专利检索	111	12.9 中文社会科学引文索引	163
9.8.2 美国专利与商标局网络 数据库	114	12.10 高校财经数据库	165
9.8.3 欧洲专利局网络检索系统	117	12.11 ProQuest Dialog(PQD) 系统	167
9.8.4 日本工业产权数字图书馆	121	12.12 EBSCO 数据库	171
第 10 章 会议信息检索	123	12.13 国道专题数据库超市 系统	175
10.1 概况	123	12.14 Emerald 数据库	177
10.2 中文会议文献检索系统 选介	124	12.15 RSC 数据库	180
10.2.1 中国重要会议论文全文 数据库	124	12.16 SciVerse ScienceDirect 数 据库	182
10.2.2 中国学术会议论文数据库	124	12.17 SpringerLink 数据库	186
10.3 美国《科学技术会议录索引》和 《社会科学及人文科学会议录 索引》	125	12.18 ACM 数据库	188
第 11 章 标准信息检索	127	12.19 ProQuest 学位论文数 据库	190
11.1 概况	127	第 13 章 电子期刊和电子图书	193
11.2 标准的种类	128	13.1 电子期刊概述	193
11.3 中国标准及其检索	130	13.2 中英文电子期刊选介	193
11.3.1 中国标准概述	130	13.2.1 中国学术期刊网络出版 总库	193
11.3.2 中国标准数据库	130	13.2.2 维普资讯数据库	198
11.3.3 万方中外标准数据库	130	13.2.3 ACS 电子期刊	200
11.3.4 中国标准服务网	131	13.2.4 John Wiley 电子期刊	203
11.4 国际标准及其检索	135	13.2.5 世界科学出版社全文电子 期刊	205
11.4.1 国际标准 ISO 及其检索	136	13.2.6 Science Online 电子期刊	208
11.4.2 国际电工标准 IEC 及其 检索	138	13.2.7 Nature 电子期刊	212
第 12 章 中英文数据库检索系统 选介	140	13.3 电子图书概述	214
12.1 中国博士学位论文数据库和中国 优秀硕士学位论文数据库	140	13.4 国内外电子图书馆选介	215
12.2 全国报刊索引数据库	140	13.4.1 EBSCO 电子图书	215
12.3 中国人民大学书报资料中心 复印报刊资料索引总汇	143	13.4.2 超星数字图书馆	218
12.4 CALIS 数据库	145	13.4.3 方正 Apabi 数字图书馆	219
		13.4.4 书生之家数字图书馆	220
		参考文献	222

第1章 概 论

21 世纪是一个全球充满竞争和挑战的时代。随着科学技术的飞速发展和全球经济一体化的推进,人类社会已经迈进了知识经济时代,知识与信息、能源、材料已经成为社会发展的三大支柱。在知识经济时代,学习将成为人们的终身需要,人们只有不断学习,不断更新知识,才能跟上时代的步伐。复杂的信息环境要求人们具有获取、加工、分析、处理信息与知识以及掌握和使用信息工具的能力,即信息素质能力。信息素质将是 21 世纪所有社会成员应具备的与读、写、算同等重要的生存能力之一,不具备信息素质的人将成为 21 世纪的新文盲。

高等院校既是出人才又是出成果的摇篮,应该培养出既有创造能力又有独立工作能力的人才,使之走向工作岗位后,善于解决问题,富有创新的勇气和技能。这就要求我们有效地利用前人已经积累的知识宝藏,把自己的研究探索工作放置于一个较高的起点之上,快、准、全地查获所需的信息,进而独立地从事学习和科研,要想做到这一点,就必须具有信息素质能力。

1.1 信息检索的意义

1. 信息检索的定义

“信息检索”一词由莫尔斯(Calvin W. Mooers)于 1949 年首次提出,并且把它定义为一种“诞生性通信形式”,“在时间上从一个时刻通往一个较晚的时刻,而在空间上可能还在同一点”,这种看法的出发点是从整体上把人类的信息传递看成是一种通信形式,即发生在当代人之间或者当代人与后代人之间的一种思想、文化、科学上的交流。而信息检索是实现信息传递的一个重要环节,它本身就意味着人类的“通信”。在某种程度上,信息检索可以和通信理论相比较。所以,从本质上来看,信息检索就是一种通信。

也有人从信息处理的角度来理解信息检索。认为信息检索主要是围绕信息检索系统如何处理信息和信息结构这一基本问题,来探讨数据库管理系统、联机检索系统等所需的硬件和软件技术,以及信息检索在各个方面的应用等问题的。这种观点把信息检索看成是计算机科学的一个分支,认为信息检索已经超出了一般意义上的文献检索的范围,还包括银行、医院、国民经济等部门中的管理信息,甚至包括经过分析和重新组织的各学科专家的知识。

还有人仅从文献信息查找的角度来表述信息检索的定义。他们研究记录下来的知识(文献)如何进行组织,才能使之有针对性地被信息用户查询和利用,认为信息检索是从大量的文献中查找与信息提问中所指定的课题(对象)有关的文献的过程。在查找过程中,计算机信息处理只是实现机械化与自动化检索系统的一种手段。根据这种观点,他们把信息检索定义为:信息检索(Information Retrieval)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和技术,所以,它的全称应该为“信息存储与检索”(Information Storage and Retrieval)。信息存储与检索是广义的信息检索,它包括信息汇集、存储与信息检索两个过程,而狭义的信息检索则仅指有序化信息的检索查找,换句话说,就是在浩瀚的信息海洋中,有的放矢,查找所需要的、有价值的或能够帮助解决问题的知识,即是从众多信息中查寻并获取所需资料信息的过程和方法,相当于人们常说的信息查询(Information Search)。

信息检索是经济生活和科学研究的一个组成部分,信息检索方法与技能在人才培养、管理和科研、生产与建设、经济技术和科学知识创新等活动中有着重要的作用。

2. 信息检索的意义

(1) 调查研究,立足创新

任何人从事某一特定领域的学术活动,或开始做一项新的科研工作,都需要花费大量的时间对有关信息进行全面调查研究,摸清国内和国外是否有人做过或者正在做同样的工作,取得了一些什么成果,尚存在什么问题,以便借鉴、改进和部署自己的工作。只有这样才能做到心中有数,才能有所发现、有所创新、有所前进。

(2) 拓宽知识面,改善知识结构

新的科学技术使人类社会生产的产业结构处在急剧的新旧蜕变之中。大批知识密集型的工业相继涌现。生产在不断发展,边缘科学大量出现,知识需要拓宽,新的知识结构是随着吸收信息量而增加的。

(3) 启迪创造性思维

知识既是过去经验的总结,又是未来的向导,古今中外一切有成就、有贡献的科学家,都是在广泛吸收前人和同代人的知识、得到启迪而取得成功的。大学生和科技人员如果能在大学期间,或者在工作中学会独立获取知识的能力,了解检索的基本知识,则是一件十分有意义的事情。这在各自的事业生涯中将会产生难以估计的效益和深远的影响。

(4) 提高自学和独立工作能力

现在高等教育的职能已经不是单纯的知识传授,而是把学生培养成信息素养人(Information Literate)。要想把他们培养成具有信息素养的人,就应该培养他们能够认识到何时需要信息,具有利用、确定、评估和有效地应用所需信息的能力。从根本意义上说,也就是培养他们知道知识是如何组织的、如何去寻找信息、如何去利用信息,使他们能为终身学习作好准备,因为他们总能寻找到为作出决策所需的信息。

(5) 减少重复劳动,提高科研成功率

科学技术史表明,科技发展的重要前提是积累、继承和借鉴前人的成果。没有继承和借鉴,就不可能有提高和创新。没有交流和综合,就没有发展。科学上的继承、借鉴、交流和综合主要是通过信息检索提供的途径来实现的。任何一个科研项目,从选题立项、实际研究到成果鉴定,每一步都离不开信息。只有充分掌握有关信息,才能避免重复,少走弯路,保证科研的高起点、高水平,缩短研究周期,获得预期效果。例如,日本高能物理研究所由于检索和利用国外的信息资料,它研制成功的第一台高能加速器的投资为40亿美元,比国外的节省二分之一。又如我国葛洲坝工程二江电站出线方案,由于信息人员及时搜集、查阅、分析了大量国内外信息资料,提出采用高压架空线路方案的建议,被采纳后,

仅投资一项就节约了 400 万元。

(6) 节省时间, 加快科研步伐

据美国和日本 20 世纪 60 年代的一个统计资料, 科学工作者在他们从事科研活动中, 用于翻阅检索信息资料的时间约占 50.9%。如果熟悉信息检索方法, 就能大大节省找资料的时间, 从而加快科研速度, 早出科研成果。例如, “汪克尔”转子发动机是德国人发明的, 日本开展研究的时间比德国晚 10 年, 但由于日本政府和有关公司全力以赴搜集、检索德国这方面的信息, 结果使得日本装有这种发动机的小汽车先于德国 10 年投入市场, 前后共为日本赢得了 20 年时间。

当今世界是一个信息开放的世界。掌握知识信息多的人对知识信息少的人就是一种挑战, 获得知识、信息快的人对慢的人就是一种压力和冲击。如何占有信息, 了解、熟悉信息检索的途径和方法, 合理地利用信息是当代大学生和科技工作者必须要解决的问题。

1.2 信息资源的类型

信息资源的主体是信息, 按照不同的标准, 可以将信息资源划分成不同的类型。按照信息资源的开发程度来区分, 信息资源可分为潜在信息资源与现实信息资源两大类。现实信息资源又可分为口语信息资源、体语信息资源、实物信息资源、文献信息资源、网络信息资源和多媒体信息资源。按照信息的应用范围及所属领域可分为科学(包括自然科学和社会科学)信息资源、技术信息资源、经济信息资源、军事信息资源、政治信息资源、生活信息资源。其中, 科学技术信息在各个方面都有着重要的作用和影响。按公开范围可分为公开信息资源和秘密信息资源。

这里只介绍现实信息资源的类型。

(1) 口语信息资源

口语信息资源, 是以口头语言, 如交谈、聊天、授课、讨论等方式获得的信息资源, 属于零次信息资源的一种。它是没有记录下来的仅靠口口相传的信息, 其特点是传递迅速、互动性强, 但稍纵即逝, 久传易出差异。代代相传的口碑、传说、口述回忆等, 虽然包含着极有价值的信息, 但是许多信息并不十分准确与可靠。因此, 通过会议、演讲、电视、广播、聊天等方式了解到的信息, 应该立即记录下来, 并加以证实。

(2) 体语信息资源

体语信息资源, 指以手势、表情、姿势, 如舞蹈、体育比赛、杂技等方式所传递的信息资源。中国人翘大拇指表示称赞, 点头表示同意。美国人耸肩表示无可奈何, 手指做成“V”, 表示“必胜”。暗送秋波、喜上眉梢都是指以眼、眉传达信息的例子。这类信息直观性强、生动丰富、印象深刻、极富感染力, 往往起到“此时无声胜有声”之效, 但此类信息的容量有限。

(3) 实物信息资源

实物信息资源, 指以实物, 如文物、产品样本、模型、碑刻、雕塑等形式表示的信息资源。例如, 秦始皇兵马俑、维纳斯雕像、人造卫星等包含着大量的信息。实物信息资源的特点是直观性强、感觉实在、信息量大, 但需要通过知识、智慧、经验和工具挖掘其隐含的大量信息。

(4) 文献信息资源

文献是记录有知识的一切载体。文献信息资源,指以文字、图形、符号、声频、视频等方式记录在各种载体上的知识和信息资源。它包括图书、连续出版物(期刊、报纸等)、学位论文、专利、标准、会议文献、政府出版物等。文献信息资源记录着无数有用的事实、数据、理论、方法、假说、经验和教训,是人类进行跨时空交流、认识和改造世界的基本工具。这类信息经过加工、整理,较为系统、准确、可靠,便于保存与利用,但也存在着信息相对滞后、部分信息尚待证实的情况。

(5) 网络信息资源

网络信息资源是指以电子资源数据的形式将文字、图像、声音、动画等多种形式的信息存放在光、磁等非印刷质的介质中,并通过网络通信、计算机终端等方式再现出来的信息资源的总和。

根据不同的标准,可将网络信息资源划分成各种类型。按照所采用的网络传输协议划分,可分成 WWW 信息资源、FTP 信息资源、Telnet 信息资源、用户服务组信息资源、Gopher 信息资源;按照信息资源的有偿性划分,可分成收费类信息资源和免费类信息资源;按照信息资源的内容划分,可分成学术研究类信息资源、教育类信息资源、政府信息资源、商业经济类信息资源、生活娱乐类信息资源和广告信息资源。

(6) 多媒体信息资源

多媒体信息资源,指将电信、电视、计算机三网相互融合,集图、文、声于一体的信息资源。它包括网上广播电视、专题论坛、网上广告等。多媒体信息打破了报刊、图书、广播、电视单项媒体的界限,形成交互式媒体信息。通过对主题、文本、模板匹配、视频检索等方式可以对其进行检索。

1.3 文献信息的类型

文献信息是当前数量最大、利用率最高的信息资源。按照各种标准,可以划分出文献的各种类型。按加工层次划分,可分为一次文献、二次文献和三次文献;按载体形式划分,可分成印刷文献、缩微文献、音像文献、机读文献。

1. 按加工层次划分

(1) 一次文献

正式出版发行的、由作者以自己的研究成果为基础创造或撰写的、未经过加工的原始文献,不论撰写时是否参考或引用了他人的资料,也不论其物质形式或出版类型如何,都称为一次文献。一般来说,期刊论文、研究报告、专利说明书、会议论文、科技报告、科技档案、技术标准以及学位论文等多属一次文献。

一次文献的主要特点包括:创造性,一次文献是科技人员在生产和研究中创造劳动的结晶,是对其创造的知识进行加工的结果;先进性,一次文献包含了在前人(包括同代人)的基础上提出的新观点、假说,或有所发明创造,达到了新的水平,或应用于新的学科和生产,反映了某一国家或国际的先进水平;价值性,一次文献能直接提供参考、借鉴和使用,可以解决生产、科研、设计、试制中的具体问题,有实际的使用价值;分散性,由于科学技术

的不断分化、综合、高速发展,致使一次文献数量大而又十分分散。这部分文献数量庞大、种类繁多,是最基本的信息源,是科技文献的主体,也是信息检索的主要对象。

(2) 二次文献

二次文献是指文献工作者将大量分散的、无序的原始文献加以筛选,留下有价值的文献,再经过加工、整理、提炼、浓缩、简化,编辑成系统的工具文献,所记录的信息称为二次信息。它包括各种目录、题录、文摘、索引等检索工具。在信息交流和传播过程中,二次文献不仅能提供大量信息,而且具有报道和便捷检索的功能。二次文献包含了大量的可供使用的信息检索工具,它可以帮助人们查找一次文献。

二次文献的主要特点包括:汇集性,二次文献将一次文献进行整序汇编,能够比较全面、系统地反映某一学科、专业或专题的文献线索,或反映某一时期、阶段的文献线索;工具性,二次文献是积累、报道和查找文献线索的一种工具,具有多种查找途径;综合性,二次文献将大量分散的文献、事实或数据进行综合、评价、筛选,以简练文字扼要叙述出来,内容概括;系统性,二次文献能及时系统地综合分析本学科、专业或专题的发展历史、现状和发展趋势,并对本领域的人物、名词术语等都有较详尽的论述,且叙述时有观点、有材料、有结论,因而系统性较强。

(3) 三次文献

三次文献是针对某一专题利用一次文献和二次文献系统地检索出一批相关文献,并对其运用科学方法和专业知识进行深入的筛选、分析、研究后重新加工、撰写的文献,如词典、手册、年鉴、百科全书、专著、教科书、论文集、述评、文献指南以及书目之书目等。一般都是参考了大量的有关文献资料,结合国内外的研究成果和发展趋势,提出具有一定高度的见解,对本领域或本研究方向具有一定的指导意义。

从一次文献到二次、三次文献,是一个由分散到集中、由片面到全面的文献加工过程。从信息工作的角度看,一次文献是检索对象,二次文献好似检索手段,三次文献是信息研究的成果。

所谓文献检索,主要指的是对二次文献和三次文献的检索。

除以上三级文献以外,还有所谓零次文献或半文献或灰色文献,是指由非正式出版物或非正式渠道传播、交流的文献,未公开于社会,只为个人或某一团体使用。例如,实验记录、设计草图、论文草稿、谈话记录、信件、未经发表的名人手迹等。

2. 按载体形式划分

(1) 印刷文献

以纸张为载体,以手写、石印、油印、铅印、胶印、影印等为手段来记录信息、知识,这是传统文献的基本形式,包括各种图书资料、连续出版物(期刊、报纸等)、学位论文、专利、标准、会议文献、政府出版物及其复印品等。就目前来说,印刷文献仍然是知识传播的主要形式。

印刷文献的主要特点是流传广泛、读取方便、不受时间和空间的限制。其缺点是信息存储密度低,占空间大、笨重,也不易保存。

1) 图书(Book)。国家标准《情报与文献工作词汇·传统文献》(GB 13143—1991)对图书的解释为:一般不少于49页并构成一个数目单元的文献。按照联合国教科文组织

(UNESCO)和国际标准化组织(ISO)规定,49 页不包括封面与扉页。48 页或少于 48 页的小书被称为小册子(Pamphlet)。图书是文献中最古老、最重要的类型。图书也称为书籍,是把记录事项按一定形式加工,归纳整理,装订成册,内容比较成熟、定型、系统、完整。科技图书大多是对已发表的科技成果、生产技术和经验或某一知识领域进行系统的论述或总结,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、重新组织而编写的。不少科技图书的内容还包括一些从未发表过的研究成果或资料。从时间上看,图书所报道的知识比期刊、论文及特种文献迟,内容一般只反映 3~5 年以前的研究水平。但科技图书中所提供的知识,一般比较全面、可靠。

图书范围较广,专著、科普读物、教科书、丛书、文集、工具书等都属于图书。

公开出版的图书都有国际标准书号(International Standard Book Number, ISBN),2007 年 1 月 1 日前,国际标准书号由 10 位数字组成,分为 4 段,中间用短线“-”隔开。第一段表示国家(0 代表英、美、加拿大、南非等英语区,1 代表其他英语区,2 代表法语区,3 代表德语区,4 代表日本,5 代表俄语区,6 代表阿拉伯语区,7 代表中国,8 代表印度,9 代表新加坡等东南亚地区),第二段是出版者号,第三段是出版社出书的序号,第四段是校验码,当校验码为“10”时用“X”代替。例如,ISBN 7-5333-0855-7。

2007 年 1 月 1 日起,实行新版 ISBN,新版 ISBN 由 13 位数字组成,分为 5 段,即在原来的 10 位数字前加上 3 位 ENA(欧洲商品编号)图书馆产品代码“978”,这样书号标识与物品标识就完整地结合在一起的。例如,ISBN 978-7-5328-5779-1。

一般说来,图书的内容成熟、系统、全面,但因出版周期长,报道速度不如期刊文献、专利文献快,不能及时提供科技信息。

2) 连续出版物(Serial Publication)。连续出版物是具有统一题名、定期或不定期以分册形式出版、有卷期或年月标识、计划无限期连续出版的文献。它包括期刊、报纸、年度出版物等。连续出版物是与图书并列的最主要的文献类型,其主要特点是内容新颖,报道及时,出版连续,信息密集,形式一致。

① 期刊(Periodical):一种定期或不定期出版的连续出版物,它具有固定的名称,有确定的出版形式,有一定出版规律,每期载有不同著者撰写的论文,按照一定的编号顺序连续出版下去。期刊上刊载的若干文献各有专题,按其内容互不相干,故期刊又名杂志。

期刊有以下几个共性:期刊刊名一般固定不变,但更名的也不少见;期刊为定期出版的连续性刊物,可分为周刊、旬刊、半月刊、月刊、双月刊、季刊、半年刊等;有连续编号,如年份、卷号、期号;不再版,一般不重印。正式出版的期刊一般都有国际标准刊号 ISSN,如 ISSN 1334-5679。

期刊文献有以下特点:期刊文献属于一次文献,它报道科研、生产中的新成就,内容新颖,反映当时的科技水平和动态,参考价值大,但其内容不如图书成熟、全面。期刊出版周期短,报道速度快。从期刊文献可及时获取有价值的科技信息,数量大、品种多、内容丰富多样。

期刊文献由于数量大、内容新、报道快,因而是科技信息的重要来源。

② 报纸(Newspaper):以报道新闻及其评论为主、出版周期比期刊更频繁、多为活页的连续出版物。它包括日报、隔日报、三日报、周报、旬报等。其特点是内容丰富、信息量大、出版迅速、发行量广、读者众多,但资料较庞杂零散,不易积累与保存,利用目录、索引

等检索工具查找与利用报纸信息,往往能起到事半功倍之效。

③ 年度出版物(Annual 或 Yearbook):汇集年内重要资料、每年出版一次的连续出版物。它包括年报、年鉴、年度进展、年度指南与手册、按年修订的百科全书等。其特点是内容较为系统,资料性强,便于查检。

3) 科技报告(Technical Report)。科技报告是科技人员围绕着某一专题从事研究,取得成果以后表述实验、研究、鉴定等工作而撰写的科学研究成果的正式报告,或者是研究过程中每个阶段进展情况的实际记录。它既不像一般的图书,又不像期刊,其主要特点是每份报告单独成一册,少的一两页,多的数百页,有机构名称,有统一编号(报告号),内容具体专深、完整可靠,反映最新研究成果。因为它是科学研究的记录和成果,因而代表着一个国家或专业的科研水平,对科学研究工作能起直接借鉴作用。科技报告按发行范围可分为绝密报告(Top Secret Report)、机密报告(Secret Report)、秘密报告(Confidential Report)、非密报告(Unclassified Report)、解密报告(Declassified Report)。解密后或根据需要,科技报告可以正式发表,但多数不公开发表,仅供有关部门使用或者参考。

科技报告基本上都是一次性文献。许多最新的研究课题与尖端学科的资料往往首先反映在科技报告中。现在全世界每年大约发表 70 多万件科技报告,占世界文献总数的 10% 左右。

科技报告大致可分为基础理论研究和生产技术两大类。若按使用类型分,有技术经济报告、生产报告、操作指标报告、设备与材料说明书报告、研究成果报告等;若按形式分,有报告书、通报、札记、论文、备忘录等;若按发表时间分,有现状报告、中间报告、总结报告等。

目前,国外常见的科技报告主要有:AD 报告——军事报告(美国武装部队技术情报局)、PB 报告——政府报告(美国政府出版局)、AEC/ERDA/DOE 报告——能源报告(美国原子能委员会)和 NASA 报告——航天报告(美国国家航空和宇宙航行局)。这就是通常所说的四大科技报告。

科技报告论述详细、系统,每份报告单独成一册出版,一般有数十页,多则上千页,因而似书不似刊,是特种文献。

4) 专利文献(Patent)。在实行专利制度的国家,政府对技术领域内的创造发明都给予法律上的保护,获得这种法律保护的创造发明就是专利。关于怎样查找专利文献,将在第 9 章详细讲解。

5) 会议文献(Conference)。会议文献是指在国际和各国国内专业学术会议上宣读的论文和交流的书面论文,是获取科技信息、了解科技发展动态的重要来源。关于怎样查找会议文献,将在第 10 章详细讲解。

6) 政府出版物(Government Publication)。政府出版物是指各国政府部门及其专门机构研究、发表的文献,包括法律、法令、议案、决议、通知、统计资料等行政性文献和科技文献等。这类文献出版发行形式多种多样,或以图书、小册子、期刊正式出版,或内部出版,或为印刷型,或为机读型。其数量庞大、内容广泛、资料可靠,是极重要的信息资源。西方国家对政府出版物极为重视,大多设有专门机构和图书馆管理机构加强对其管理和利用。借助于政府出版物,可以了解某一国家的科技政策、经济政策,且对于了解其科技活动、科技成果也具有一定的参考价值。

政府出版物的最大特点是重复。很多文件在政府颁布之前,部分内容已由一些相关单位出版过,有一定的重复,行政性文件的重复出版更为多见。

7) 标准文献(Standard Literature)。标准文献有狭义和广义之分。狭义指按规定程序制定,经公认权威机构批准的一整套在特定范围内必须执行的规格、规则、技术要求等规范性文献。广义指与标准化工作有关的一切文献,包括标准形成过程中的各种档案、宣传推广标准的手册及其出版物、揭示报道标准文献信息的目录、索引等。标准文献的检索将在第 11 章详细讲解。

8) 学位论文(Dissertation)。学位论文是高等院校、科研单位的研究生、本科生为了取得学位而撰写的论文。在我国,学位论文分学士、硕士和博士论文。各级学位论文都是在有专长的导师指导下研究成功后撰写的,并经过答辩和评审,具有一定质量。论文的研究主题,大多数是某学科的新课题,具有一定独创性。学位论文对问题的探讨比较专深,特别是博士论文,具有较高的学术水平。此外,学位论文叙述也比较系统和详细。因此,学位论文有一定的信息参考价值。

大多数学位论文属于非卖品,除少数印成单行本或在期刊上、学术会议上发表外,一般都不出版发行,因而份数有限,馆藏分散,搜集困难。不过,现在有些机构建立了博硕士学位论文数据库,为查阅博硕士学位论文提供了方便。

9) 产品样本(Trade Catalogue)。产品样本又叫产品目录、产品说明书或产品资料,是对已经投入生产的产品作介绍的资料。换句话说,产品样本是对定型产品的性能、构造、原理、用途、使用方法和操作规程、产品规格等所作的说明。它图文并茂,形象直观,出版发行迅速,更新较快,多数由厂商赠送。从产品样本中可以获得关于产品结构的详细说明,并且由于它代表已投产的产品,在技术上比较成熟,数据比较可靠,有较多的外观照片和结构图,直观性强。查阅和分析国内外有关产品样本,有助于了解国内外产品的生产水平、工艺水平、技术水平及其有关技术的演变过程和发展动向,可以获得设计、制造、使用中所需要的数据和方法,也可以判断某种产品的价值,为我国引进设备提供信息。在使用产品样本时,也要看到它的局限性,即真正的技术关键和所谓的核心数据在产品样本中一般是不介绍的,使用时要加以识别,应同其他类型文献,如专利、期刊论文、标准等配合使用,以攻破其核心技术难点部分。

10) 科技档案(Archives)。科技档案是围绕着科研、生产而形成的技术文件,诸如科技规划或计划、方案、任务书、协议书、技术指标、操作规程、图表、技术鉴定、经济核算等。

科技档案有以下特征。

① 专业性:科技档案是在一定专业范围内科技活动的真实记录,所以大多数科技档案按专业分类。

② 成套性:一件科技档案包括科技活动中准备、实验、总结各阶段形成的全部技术文件,它们往往被成套地收藏下来。

③ 准确性:科技档案不仅反映科技活动过去的历史,更应如实地、准确地反映科技发展的现状。和文书档案不同,经过一定审批手续,科技档案允许修改,使之和科技发展水平相适应。

④ 通用性:它是指科技档案可以提供给他人充分利用。据统计,在设计部门,科技人员重复利用现存的设计成果往往达到 50%~70%。引进国外先进技术,在某种意义上