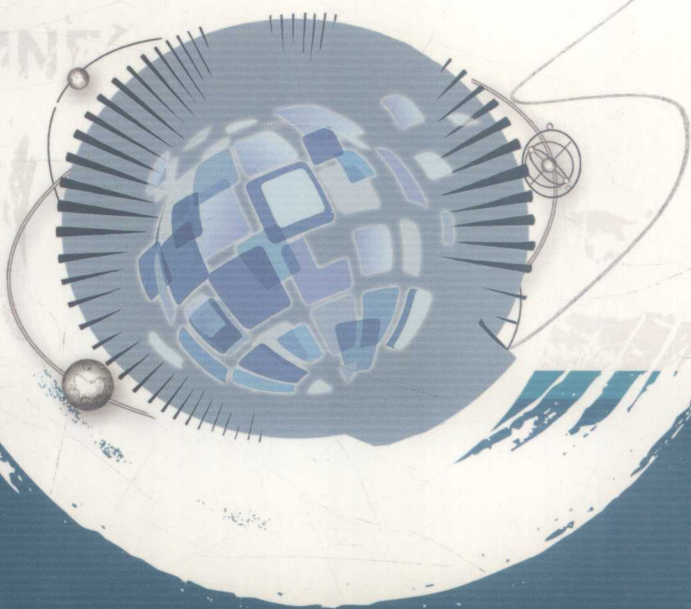




21世纪普通高等学校信息素质教育系列规划教材

现代信息检索与利用

樊 瑜 / 主编



华中科技大学出版社

<http://www.hustp.com>

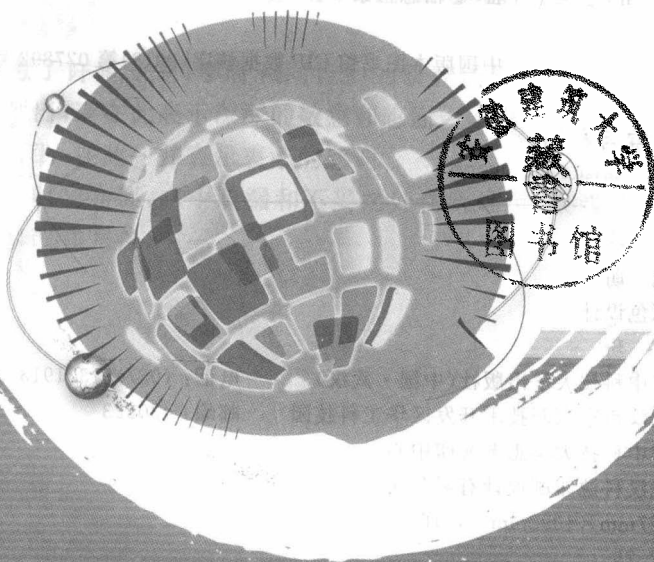


21世纪普通高等学校信息素质教育系列规划教材

现代信息检索与利用

◎ 主 编 樊 瑜

◎ 副主编 武宗锋 魏现辉 闫晓妍



华中科技大学出版社

<http://www.hustp.com>

中国·武汉

内 容 提 要

本书注重数据库和网络信息的检索,满足了在校生的需求,详细论述了信息资源各个类型的检索方法和途径,叙述清晰,条理性和逻辑性强,便于读者阅读和自学;同时,本书具有较强的系统性、实用性和可操作性,读者按照书中的操作实例,对照操作,便可以完成检索步骤,得到所需的检索结果。

本书共分十五章。全书基本按照文献出版类型进行编排,注重数据库和网络信息的检索。第一章至第三章为原理篇;第四章为网络检索工具原理篇;第五章至第七章为图书类型文献的检索;第八章为连续出版物类型文献的检索;第九章至第十一章为特种文献类型的检索;第十二章为外文文献的检索;第十三章为高校在校学生常用辅助学习资源的检索和使用;第十四章为解决本单位资源不足而通过资源共享联合又能实现的文献传递的使用;第十五章为信息检索的应用篇,指导在校学生毕业论文的写作和科研工作者专业论文的写作。

本书可供高等院校在校师生教学使用,也可供科研人员参考,对使用信息资源数据库和网络信息检索有较高的参考价值。

图书在版编目(CIP)数据

现代信息检索与利用/樊瑜主编. —武汉:华中科技大学出版社,2018.2
ISBN 978-7-5680-3811-9

I. ①现… II. ①樊… III. ①信息检索 IV. ①G254.9

中国版本图书馆 CIP 数据核字(2018)第 027892 号

现代信息检索与利用

Xiandai Xinxi Jiansuo yu Liyong

樊 瑜 主 编

策划编辑:袁 冲

责任编辑:沈 萌

封面设计:原色设计

责任监印:朱 玟

出版发行:华中科技大学出版社(中国·武汉)
武汉市东湖新技术开发区华工科技园

电话:(027)81321913

邮编:430223

录 排:华中科技大学惠友文印中心

印 刷:武汉科源印刷设计有限公司

开 本:787mm×1092mm 1/16

印 张:17.75

字 数:439 千字

版 次:2018 年 2 月第 1 版第 1 次印刷

定 价:39.00 元



本书若有印装质量问题,请向出版社营销中心调换
全国免费服务热线:400-6679-118 竭诚为您服务
版权所有 侵权必究

前 言

我们已迈入知识经济时代,信息资源作为一种取之不尽、用之不竭、通过加工处理可再生的资源,是继材料、能源之后出现的第三种重要资源,已经成为支撑当今社会发展的三大支柱之一。一个国家的科技进步和社会发展,越来越多地取决于对信息的利用与开发。谁能充分、有效地利用信息资源,谁就能抢占科学技术发展的制高点。如何获取和利用信息资源成为信息社会人们学习的前提和基础。信息素养已成为国民综合素养的一项重要指标,为了提高大学生的综合素养,国内外高校纷纷将信息素养教育作为培养人才的重要内容。

信息检索是信息素养教育的主体,是实施信息素养教育的必修课程,是培养学生信息意识、独立学习和终身学习能力,使其掌握从海量信息中获取所需文献信息的方法和工具,目标是增强学生的信息利用和知识创新能力。

本书编者在总结近 20 年教学经验的基础上,将理论与实践紧密结合,根据具体数据库的检索,强化了信息检索技术方面的内容,突出了实用性。全书按照文献类型进行组织排序,突出实用性、直观性和新颖性,特别是将文献传递、随书光盘和行业信息的检索和利用等内容增加在里面,丰富了传统信息检索的内容,适合不同专业在校学生和技术人员使用。

本书由樊瑜任主编,负责策划、定稿工作。全书共分为十五章。第一章、第二章、第五章、第七章由魏现辉编写;第三章、第六章、第九章、第十章、第十一章、第十四章由樊瑜编写;第八章、第十三章、第十五章由武宗锋编写;第四章、第十二章由闫晓妍编写。全书由吴少杰负责主审。

本书在编写过程中,参考了许多专家、学者和同行们的研究成果、论著、论文及相关材料,有的已经在参考文献中列出,谨在此表示深深的敬意和感谢。本书的编写还得到了崔慕岳、张怀涛等前辈的指导和帮助,在此一并表示衷心的感谢。

由于编者水平有限,加之时间仓促,书中疏漏和不妥之处在所难免,敬请学术界同仁和读者批评指正,以便进一步修改和完善。

樊 瑜

2017 年 12 月

目 录

第1章 绪论	(1)
1.1 信息素养与信息意识	(1)
1.2 信息检索的意义与大学生信息素质教育	(3)
1.3 信息及相关术语	(4)
1.4 信息资源概述	(6)
第2章 信息检索基本知识	(10)
2.1 信息检索的概念和类型	(10)
2.2 信息检索工具	(11)
2.3 信息检索语言	(15)
2.4 信息检索的一般方法和途径	(22)
第3章 计算机信息检索	(25)
3.1 计算机信息检索的基本原理	(25)
3.2 计算机信息检索系统	(28)
3.3 计算机信息检索基本技术	(29)
3.4 计算机信息检索的策略与技巧	(33)
3.5 计算机信息检索效果的评价	(38)
第4章 网络信息检索	(41)
4.1 网络信息检索概述	(41)
4.2 搜索引擎	(42)
4.3 百度搜索引擎	(44)
4.4 百度地图搜索	(48)
第5章 OPAC检索	(54)
5.1 OPAC简介	(54)
5.2 馆藏目录检索	(56)
5.3 联合目录检索	(63)
第6章 随书光盘数据库检索	(67)
6.1 博云非书资料管理系统	(67)
6.2 OPAC光盘检索	(72)
6.3 畅想之星光盘中心	(73)
第7章 电子图书检索	(80)
7.1 电子图书概述	(80)
7.2 超星数字图书馆	(82)
7.3 方正 Apabi 数字图书馆	(88)
7.4 读秀中文学术搜索	(91)



第 8 章 中文连续出版物检索	(95)
8.1 中国知网中国学术期刊(网络版)	(95)
8.2 维普资讯中文期刊服务平台 7.0	(98)
8.3 万方中国学术期刊数据库(新版)	(101)
第 9 章 标准文献检索	(105)
9.1 标准文献概述	(105)
9.2 中国标准文献检索	(109)
9.3 国际标准及其检索	(128)
9.4 国外先进标准及其检索	(135)
第 10 章 专利文献检索	(141)
10.1 专利概述	(141)
10.2 专利文献	(144)
10.3 专利文献分类	(150)
10.4 专利文献编号	(152)
10.5 专利文献检索概述	(160)
10.6 专利文献的网络检索	(161)
第 11 章 会议文献、学位论文和科技报告检索	(172)
11.1 会议文献检索	(172)
11.2 学位论文检索	(181)
11.3 科技报告检索	(191)
第 12 章 国外文献信息检索	(200)
12.1 《工程索引》及其数据库	(200)
12.2 《科学引文索引》及其数据库	(206)
12.3 SpringerLink 数据库	(210)
12.4 CALIS 外文期刊目次数据库	(215)
第 13 章 高校在校学生常用数据库检索	(220)
13.1 银符考试题库 B12 数据库	(220)
13.2 新东方多媒体学习库	(234)
13.3 新东方掌学平台	(236)
13.4 爱迪科森网上报告厅	(238)
13.5 超星名师讲坛	(241)
第 14 章 文献传递	(243)
14.1 中国高等教育文献保障系统文献传递	(243)
14.2 开世览文文献传递	(246)
14.3 超星读秀学术搜索文献传递	(250)
14.4 “百链”文献传递	(257)
第 15 章 学术论文写作	(261)
15.1 学术论文概述	(261)
15.2 高校师生常见学术论文类型简述	(262)
15.3 学术论文的写作方法	(265)
参考文献	(276)

第1章 绪 论

1.1 信息素养与信息意识

1.1.1 信息素养

信息素养(也称信息素质或信息文化)的本质是全球信息化需要人们具备的一种基本能力,即能有效地选择、寻找及评估传统或网上资源的技巧。信息素养这一概念最早是由美国信息产业协会主席保罗·泽考斯基(Paul Zurkowski)于1974年提出的,即“人们利用多种信息工具使问题得到解答的技术和技能”。简单的定义来自1989年美国图书馆学会,包括:能够判断什么时候需要信息,并且懂得如何去获取信息,如何去评价和有效利用信息。

信息素养是一种基本能力。信息素养是一种对信息社会的适应能力。美国教育技术CEO论坛2001年第4季度报告提出21世纪的能力素质,包括基本学习技能(读、写、算)、信息素养、创新思维能力、人际交往与合作精神、实践能力。信息素养是其中一个方面,它涉及信息的意识、信息的能力和信息的应用。

信息素养是一种综合能力。信息素养涉及各方面的知识,是一个特殊的、涵盖面很宽的能力,它包含人文的、技术的、经济的、法律的诸多因素,和许多学科有着紧密的联系。信息技术支持信息素养,通晓信息技术强调对技术的理解、认识和使用技能;而信息素养的重点是内容、传播、分析,包括信息检索以及评价,涉及更宽的方面。它是一种了解、搜集、评估和利用信息的知识结构,需要通过熟练的信息技术和完善的调查方法以鉴别和推理来完成。信息素养是一种信息能力,信息技术是它的一种工具。

1998年,美国图书馆协会和教育传播协会制定了学生学习的九大信息素养标准,概括了信息素养的具体内容。

- (1)具有信息素养的学生能够有效地和高效地获取信息。
- (2)具有信息素养的学生能够熟练地和批判地评价信息。
- (3)具有信息素养的学生能够精确地、创造性地使用信息。
- (4)作为一个独立学习者的学生具有信息素养,并能探求与个人兴趣有关的信息。
- (5)作为一个独立学习者的学生具有信息素养,并能欣赏作品和其他对信息进行创造性表达的内容。
- (6)作为一个独立学习者的学生具有信息素养,并能力争在信息查询和知识创新中做得最好。
- (7)对学习社区和社会有积极贡献的学生具有信息素养,并能认识信息对民主化社会的重要性。



(8)对学习社区和社会有积极贡献的学生具有信息素养,并能实行与信息和技术相关的符合伦理道德的行为。

(9)对学习社区和社会有积极贡献的学生具有信息素养,并能积极参与小组的活动探求和创建信息。

参考上述标准,结合国内高等教育的实际情况,大学生的信息素养培养主要针对以下五个方面的内容。

(1)热爱生活,有获取新信息的意愿,能够主动地从生活实践中不断地查找、探究新信息。

(2)具有基本的科学和文化常识,能够较为自如地对获得的信息进行辨别和分析,正确地加以评估。

(3)可灵活地支配信息,较好地掌握选择信息、拒绝信息的技能。

(4)能够有效地利用信息、表达个人的思想和观念,并乐意与他人分享不同的见解或信息。

(5)无论面对何种情境,能够充满自信地运用各类信息解决问题,有较强的创新意识和进取精神。

1.1.2 信息意识

信息意识是信息素养的前提,是人对信息的敏感程度,是人对信息敏锐的感受力、持久的注意力和对信息价值的洞察力、判断力等。它决定着人们获取、判断和利用信息的自觉程度。人们对自己感兴趣或与自身利益相关的信息,感受力通常非常敏锐,注意力也相对持久,对信息价值的判断力和洞察力也更为强烈,这样才能够从大量的事物和社会现象中敏锐地发现并获取有价值的信息。反之,对与自身无关或相关性弱的信息则缺乏相应的感知和获取能力。

信息意识包括信息主体意识、信息获取意识、信息传播意识、信息更新意识、信息安全意识等。

当今社会,随着计算机技术和通信技术的飞速发展以及互联网的出现和广泛普及,信息已经无处不在。针对这一状况,要想快速、准确地获取信息,必须提高信息意识。

提高信息意识是提高掌握知识能力的要求。当代社会学家和经济学家一致认为,在知识经济社会,人类面临的主要问题就是知识问题,只有知识才是唯一可持续发展的资源,其他资源都是依附于知识的配置资源,而知识问题的解决,最终落脚点在提高社会整体信息交流的效率与效用。我们生活在由数据、信息、知识构成的信息流中,面临的主要任务是如何实现由数据到信息再到知识的转换。因此,在信息时代,知识就是利用信息的能力。

提高信息意识是开放性学习和终身教育的要求。知识的快速增长使人们掌握和利用所需要的信息和知识变得越来越困难。因此,开放性学习、终身学习就显得至关重要。对于大学生来说,提高学习的能力比学习知识本身更为重要。学生们要学会怎样获取专业信息、怎样对获得的大量信息进行筛选、如何去选择相关的信息及运用这些信息去构筑自己的知识体系,培养从记忆信息到应用信息和创新信息的能力。

提高信息意识是进入社会的要求。随着人类社会步入全新的信息时代,科技的进步和社会经济的发展对信息资源、信息技术的依赖性越来越大。在这样的社会大环境中,如果一



个人没有一定的信息意识,不具备信息获取、分析和利用的基本能力,那么他就不具有竞争力,必将被社会所淘汰。

因此,提高人们的信息意识日益成为社会高度关注的重要问题,同时也是当今国内外高等教育中十分重视的热点问题。信息时代呼唤科学的信息素质教育,已成为全球教育界的共识。培养现代信息意识、提高自身的信息素养已经成为当代每一位大学生的当务之急。

1.2 信息检索的意义与大学生信息素质教育

1.2.1 信息检索的意义

在计算机、网络快速发展,互联网日益普及的今天,我们置身于信息与知识的海洋中,要想及时准备、全面地查找所需的信息资源,不学习和掌握信息检索的知识和方法是行不通的。只有随时随地发现和获取身边无处不在的信息,获得新知识,才不会被瞬息万变的社会所淘汰。而要想高效、快速地掌握最新信息,获取最新知识,信息检索就成为满足信息需求的必要途径。概括来说,信息检索的重要意义与作用主要体现在以下几个方面。

(1) 信息检索是获取知识的有效途径。

在当今文献知识急剧增长的信息时代,人类获取信息面临三大挑战:一是知识信息的无限性对人们有限的阅读时间的挑战;二是迅速增长的信息量对人类接受能力的挑战;三是大量新知识对人们理解能力的挑战。据测算,人类知识总量在19世纪每50年增加一倍,20世纪初每10年增加一倍,20世纪70年代每5年增加一倍,20世纪80年代几乎每3年增加一倍,进入21世纪,知识总量更是以指数级递增。这就要求人们必须通过不断学习,获取和积累新知识才能适应社会发展的需求。因此,如何以最少的精力、最短的时间充分占有文献信息,成为人们亟待解决的实际问题,而信息检索正是解决这一问题的最好途径,它可以帮助人们快速、准确、全面地获取所需信息,最大限度地节省查找时间,使文献信息得以充分利用。因此,掌握信息检索的方法和技巧,是获取知识和更新知识的重要手段,是做到无师自通、不断进取的主要途径。

(2) 信息检索能提高科研效率,避免重复劳动。

科学发展的历史证明,积累、继承和借鉴前人或他人的研究成果是科学发展的重要前提。在当代科学研究和技术开发过程中,一个研究人员查找资料 and 了解同行工作进展信息的时间占研究工作时间的一半以上。据美国国家基金会在化学工业部的调查统计表明,研究人员的全部工作时间分配是:计划与思考占7.7%,信息收集占50.9%,试验研究占32.1%,数据处理占9.3%。正如牛顿所说:“假如说我比前人看得远一点,那是因为我站在巨人肩膀上的缘故。”科学研究具有继承和创造两重性,这就要求科研人员在探索未知或从事研究工作之前,尽可能地占有与之相关的信息,即利用信息检索的方法充分了解国内外相关研究的成果及研究现状和发展动向,选择研究的最佳切入点。只有这样才能正确地制订研究方案,避免重复研究,尽可能降低获取信息和知识的成本。因此,信息检索是科学研究活动中提高效率、避免重复劳动必不可少的前期工作。

(3) 提供科学方法,为管理者提供决策参考。

科技与经济的管理决策同样离不开信息。任何个人、企业乃至国家,要想在竞争中



立足,都必须掌握足够、准确的信息,并利用它们进行科学决策。信息获取成功的基础则是通过科学合理的信息检索方法获取大量有用的信息。信息检索为人们提供了一套完整的开发、利用信息资源的方法,包括信息检索工具的选择、信息检索策略的制订、信息检索手段的选择等。在激烈的市场竞争中,管理者必须时刻关注竞争对手的动向,力求扬长避短,确立自己的竞争优势。“优胜劣汰、适者生存”是市场竞争的自然规律。要想在激烈的市场竞争中立于不败之地,首先要有科学的决策。企业在市场中要不断开发新产品,选择投资项目,确定营销策略,这一切都离不开准确及时的市场信息。因此可以说,信息竞争是企业成败的关键,它是市场导向的风向标,是现代企业生存发展的战略武器和重要保障。

1.2.2 大学生信息素质教育

近年来,我国已逐步认识到在高校开展信息素质教育的重要性。2002年1月,教育部高校图书馆工作指导委员会召开了“全国高校信息素质教育学术研讨会”,这次会议被认为是我国高校图书馆从用户教育阶段过渡到信息素质教育阶段的标志。从2004年起,在我国图书馆界召开的多个学术会议中,信息素质教育都是其中的一个重要议题。

信息素质教育主要包含以下三个方面。

(1) 信息认知能力的培养。

通过充分了解信息、信息资源、信息技术、信息产业等基本概念,认清信息化的发展趋势和影响;理解信息的社会功能,理解信息资源、信息技术对社会变革、经济增长、文化发展所起到的作用;充分认识到信息在当今社会的重要性。从而在实际工作、学习和生活中充分认识到自身的信息需求,及时地获取信息,并利用相关信息来解决实际问题。

(2) 信息知识及能力的提高。

信息知识包含对信息的分布、组织加工、传播、检索利用等各个环节的原理、现状、规律的认识。只有不断更新信息知识、积极尝试利用检索技能解决实际问题,才可能具备现代社会所要求的信息能力,包括信息检索能力、信息识别能力、信息处理能力、信息组织能力、信息评价能力等。

(3) 信息道德及法制的教育。

信息道德是指在整个信息活动中的道德规范,它是调节信息创造者、信息服务者、信息使用者之间相互关系的规范。每个社会成员在信息活动中都应该遵守法律法规,尊重他人的学术成果,尊重知识产权,合理利用信息,自觉抵制信息违法行为。

1.3 信息及相关术语

当今时代,信息无处不在,网络、电视、广播、报纸等各种媒体每天都向公众传播着各种各样的信息。处在这样的环境中,我们每个人都能接收到各种类型的信息,并自觉或不自觉地参与信息的传递。那么,究竟什么是信息?信息、知识、情报之间的关系怎样?这就是我们本节要学习的问题。

1.3.1 信息(Information)

“信息”是一个既古老又年轻的概念。早在两千多年前的我国西汉时期,即有“信”字的



出现。唐朝诗人杜牧在《寄远》一诗中喟叹“塞外音书无信息，道傍车马起尘埃”；南唐诗人李中在《暮春怀故人》中也写下“梦断美人沉信息，目穿长路倚楼台”的佳句。很明显，信息在此处指“音讯”“消息”等。

信息作为一个科学术语被提出和使用，可追溯到1928年哈特莱(Hartly R. V.)在《信息传输》一文中的描述。他认为：“信息是指有新内容、新知识的消息。”1948年，美国数学家、信息论创始人申农(Shannon C. E.)在《通信的数学理论》一文中指出：“信息是用以消除随机不确定性的东西。”也就是说，信息接收方在未收到信息前不知道信息产生方发出什么信息，只有在收到消息后才能消除信源的不确定性。控制论的创始人维纳(Wiener N.)在其专著《控制论——动物和机器中的通信和控制问题》中，阐述信息是“我们在适应外部世界、控制外部世界的过程中，同外部世界交换内容的名称”。显然，他把人与外部环境交换信息的过程看成一种广义的通信过程。1956年，英国学者Ashby E.提出：“信息是集合的变异度，信息的本质在于事物本身具有变异度。”1975年，意大利学者Longo G.在《信息论：心得趋势与未决问题》中指出：“信息是反映事物构成、关系和差别的东西，它包含在事物的差异之中，而不在事物的本身。”2011年又出现了“信息是反映事件的现象、确定性、属性、构成、关系和差别等的内容”的说法。

可见，随着人们对信息本质认识的逐步加深，信息的概念也在不断地拓展和丰富。迄今为止，信息仍然没有权威的、统一的定义。本书从信息检索与利用的角度出发，给信息一个比较广泛的、容易理解的定义：信息是一切事物存在方式和运动状态的客观反映。

1.3.2 知识(Knowledge)

知识是人们在改造客观世界的实践中所获得的认识和经验的总结，即人类对客观世界的认识。人们在认识世界和改造世界的过程中，获得了大量事物运动变化的规律，即感性认识和经验，然后通过大脑的组织和加工处理，形成理性认识，即知识。因此，人类不仅要通过信息感知世界，还要对获得的信息进行整理和加工，形成知识并进一步认识和改造世界。知识可以分为两种，即主观知识和客观知识。主观知识主要指人们头脑中的经验、观点、思想等。在文字出现以前，知识主要以主观知识的形式出现。随着人类文明的发展，文字出现以后，才产生了以文字、图形为记录载体的客观知识，从而实现了知识信息的跨时空交流、传递与利用。可见，知识是信息的一部分。

1.3.3 情报(Intelligence)

情报是在人类社会传递中可以发挥效益的关于特定事实的信息和知识。情报一般由内容、符号、载体构成。根据所属的范围可以分为专业情报、管理情报、军事情报等；根据情报的重要性可分为重要情报、一般情报等。情报的概念在不同历史时期有着不同的含义：起初，情报是特指战时关于敌情的报告；20世纪70年代，情报是指决策、部署、规划、行动所需要的能指引方向的知识与智慧；20世纪80年代，情报又被认为是获得其他方的有关情况以及对其分析研究的结果。无论情报的内容与形式如何变化，情报都具有知识性、传递性、效用性。

1.3.4 文献(Document)

中华人民共和国国家标准GB 3792.1—1983《文献著录总则》中对文献的定义是：“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/DIS 5217)对文献的定义是:“存贮、检索、利用或传递记录信息过程中,可作为一个单元处理的、在载体上或依附载体而存贮有信息或数据的载体。”

从上述定义可以看出,构成文献,必须具备四个要素,即文献信息、文献载体、符号系统和记录方式。文献信息是文献的内容,它属于信息,具有信息的一切性质和功能;文献载体是文献的外部形式,即材质,它在人类文明进程中一直不断地演变,从早期的石块、兽骨、金属、纸张到现代的胶片、磁带、光盘等;符号系统是信息的携带者,包括图形、文字、公式、图表、编码等;记录方式是代表文献的符号通过特定的人工记录方式进入载体的方法和过程。

文献对人类文明的进步具有重要意义,主要体现在两个方面:一是知识的积累、传播和创造;二是对研究成果的确认和评价。文献是重要的信息资源,它能帮助人们克服时间与空间的限制,记录、贮存和传递人类已有的知识与经验,所载信息可供无数人同时异地反复使用,还可以通过复制广泛传播,从而给人类社会带来巨大的社会效益和经济效益,推动人类社会不断发展和进步。

1.3.5 知识、情报、文献与信息的关系

知识、情报、文献和信息之间存在着内在的必然联系。信息是物质的属性,是物质的一种存在形式,它以物质的属性或运动状态为内容,并且总是借助于一定的物质载体存储和传递。知识的产生离不开信息和信息的传递,知识包含在信息之中。情报是特殊的信息,是在一定时间内为达到一定目的而传递给特定对象的、有用的知识或信息。在人类漫长的生产实践、科学试验以及社会实践过程中,各种现象及经验等相关知识信息不断地积累,为了便于记忆、交流和传播,于是产生了文献。

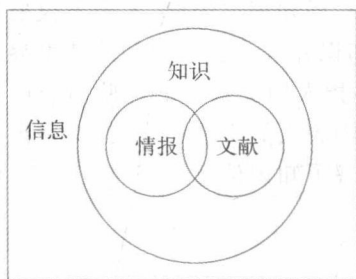


图 1-1 信息、知识、情报和文献之间的关系

综上所述,信息是起源和基础,它包含了知识和情报。文献则是信息、知识、情报的存储载体和传播工具,是重要的知识源。信息可以成为情报,但是一般要经过选择、分析、研究等加工过程。知识是经过总结、提炼和系统化后的信息。信息、知识、情报的主要部分被包含在文献之中,但远非全部。目前,学术界比较一致的看法是:信息>知识>情报,文献和情报的关系非常密切,并且有所交叉。信息、知识、情报和文献之间的关系如图 1-1 所示。

1.4 信息资源概述

信息资源是信息与资源两个概念整合衍生出的新概念。要了解信息资源,我们先从信息资源的概念入手。

1.4.1 信息资源的概念

信息资源的概念有狭义和广义之分。

狭义上认为信息资源是指人类社会经济活动中经过加工处理有序化并大量积累起来的

有用信息的集合。

广义上认为信息资源是指人类社会信息活动中积累起来的信息、信息生产者、信息技术等信息活动要素的集合。信息资源包括下述几个部分:①人类社会经济活动中经过加工处理有序化并大量积累起来的信息、信息生产者;②为某种目的而生产信息的信息生产者的集合;③加工、处理和传递信息的信息技术的集合;④其他信息活动要素(如信息设备、设施、信息活动经费等)的集合。

一般情况下,我们从狭义上来理解信息资源,即信息资源是指人们通过一系列的认识和创造过程,采用符号形式储存在一定载体(包括人的大脑)之上的、可供利用的全部信息。

信息是普遍存在的,但并非所有的信息都是信息资源,只有经过人类开发与重新组织后有利用价值的信息才称为信息资源,即信息资源是信息世界中对人类有价值的那一部分信息,是附加了人类劳动、可供人们利用的信息。因此,构成信息资源的基本要素是信息、人、符号、载体。信息是组成信息资源的原料,人是信息资源的生产者和利用者,符号是记录和描述信息资源的手段,载体是存储和利用信息资源的物质形式。信息资源与其他资源相比,具有可再生性和可共享性的特点。可再生性是指它不同于普通的自然资源,它可以被反复使用而不失去其利用价值。对它的开发利用越深入,它的内容越丰富,利用价值越高。可共享性是指它能被不同用户分享而不失去其原有信息量。

随着信息社会的到来,人们越来越清晰地认识到信息资源的重要性。当前,人们从事的一切社会活动都必须从利用信息资源入手,在未来的竞争中,那些占有较多信息资源的国家、机构和个人将具有更大的竞争优势。信息资源地位的不断上升,促进了以计算机和网络为核心的信息技术的进步,先进的信息技术又反过来推动了信息资源的开发和利用。

1.4.2 信息资源的类型

要开发和利用信息资源,就必须首先了解信息资源的类型。按照不同的分类标准,信息资源类型的划分方法也不同。

按信息资源的存在状态可将其分为潜在信息资源和现实信息资源。潜在信息资源是指人类在认识和思维创造过程中,存储在大脑中的信息,只能为本人或有限的人所利用,是一种有限再生的信息资源。现实信息资源是指人脑中的信息通过特定的符号和载体表述后,可以在特定条件下广泛地传递并连续往复地为人类所使用,是一种无限再生的信息资源。

显然,现实信息资源是我们当前研究、开发和利用的重点。现实信息资源依据其载体可分为载体信息资源、实物信息资源、文献信息资源和网络信息资源。

(1)载体信息资源。载体信息资源是指以人体为载体并能被他人识别的信息资源,包括口语信息资源和体语信息资源,如讲话、授课、讨论、表情、姿态等都属于载体信息资源的范畴。

(2)实物信息资源。实物信息资源是指以实物为载体的信息资源,即人类通过创造性劳动以实物形式表述的信息资源,通常以样品、模型、雕塑等实物进行展示与交流。

(3)文献信息资源。文献信息资源是指以文献为载体的信息资源。它又可以根据载体、出版形式和信息加工程度进行划分。

①按照载体划分。

a. 印刷型信息资源。这是一种传统的最为常见的信息资源,是指通过油印、铅印、胶印



等各种印刷手段将信息记录在纸张或其他类纸材质上的信息资源。其特点是便于阅读和传播,但占用空间大,不易整理和保存。

b. 缩微型信息资源。通过利用光学技术将信息记录在感光材料上的信息资源,包括缩微胶卷和缩微平片。其特点是存储密度高、易保存,但需要专门的设备才可使用。

c. 声像型信息资源。通过专门的设备,使用声、光、电技术将信息以声音、图像等形式记录下来的信息资源,包括唱片、录音带、幻灯片等。其特点是直观生动,但需要专门的设备。

d. 电子型信息资源。通过编码技术将信息转换为计算机可识别的代码,并存储在磁盘、光盘上的信息资源。它具有存储量大、存取速度快、占用空间小、远距离瞬时传输等特点。这也是目前发展速度最快、使用最广泛的一种信息资源。

②按照出版形式划分。

a. 图书。图书是指将文字、图形或其他符号书写或印刷于纸张上的、具有完整装帧形式的非连续出版物。图书一般内容系统全面、主题突出、观点成熟,但编辑出版周期较长,报道速度相对较慢。按用途可分为阅读型图书和工具型图书,按版本形式可分为单卷书、丛书、专著、参考书等。图书是传播知识、教育和培养人才的主要工具。

b. 期刊。期刊是一种有比较固定的名称、有固定出版规律的连续出版物。它出版周期短、报道速度快、内容新颖丰富,能及时反映当代社会生活和科技的发展水平与动向。

c. 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连续出版物,通常每天或定期以散页形式出版。它报道的内容非常广泛,和人们的生活息息相关,是人们生活中经常接触到的信息资源。主要特点:出版周期很短,能以最快的速度报道国内外发生的新事件和科技的最新研究成果。报纸的信息具有很强的实效性,信息量大,但查找不够方便。

d. 专利说明书。也称专利文献,是特指专利申请人向专利主管部门呈交的有关发明创造的详细技术说明书,是具有自是产权特性的信息资源,主要包括经审批授权的专利说明书和未经审批的专利申请公开说明书,一般由专利主管部门出版发行。专利说明书涉及的技术内容广泛,从国防尖端技术到普通的工程技术以及日常生活用品,无所不包,具有融技术信息、经济信息、法律信息于一体的特点。根据世界知识产权组织的统计,全世界每年的发明创造成果信息 90%~95% 都能在专利说明书中查到,并且大部分发明创造只通过专利说明书公开。

e. 科技报告。科技报告是描述一项研究的进展或取得的成果,或一项技术研制试验和评价结果的一种文本。科技报告每份单独成册,有专门编号,用以识别报告类型及其主持机构。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,并且有一定的保密性,相当一部分科技报告不公开发行,是获取最新技术研究成果信息的重要信息资源。世界上比较著名的科技报告是美国的四大报告:AD(军用工程)、PB(民用工程)、DOE(能源工程)、NASA(航空航天工程)。

f. 技术标准。也称标准文献,是描述有关产品的工程质量、规格、工艺流程及其测试方法等的技术文件。它是一种经权威机构批准的规章性文献。技术标准有单行本和汇编本两种出版发行方式。它具有计划性、协调性、法律约束性等特点。在一定区域内施行统一的标准可促使产品规格化、系列化和通用化,对提高生产水平、产品质量,推广应用研究成果,推动科技发展等有着十分重要的作用。

g. 政府出版物。政府出版物是各国政府部门及其所属机构编辑出版的具有官方性质的

文献。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。政府出版物在出版前后,往往以其他形式发表,内容有时与其他类型的文献(如科技报告)有所重复。它是政府用以发布政令和体现其思想、意志、行为的物质载体,同时也是使政府的思想、意志、行为产生社会效应的主要传播媒介。

③按照信息加工程度划分。

a. 零次文献信息资源。未经正式出版发行的最原始的记录,如书信、手稿、笔记、实验记录等。其特点是内容新颖、具有原始性,但不成熟、分散,难于检索。

b. 一次文献信息资源。以作者本人的研究工作成果为依据撰写的,已公开发行的专著、学术论文、专利说明书、科技报告等。因此,一次文献信息资源包含了新观点、新技术、新发明、新成果,具有创造性的特点,有直接参考、借鉴和使用的价值,是人们检索和利用的主要对象。

c. 二次文献信息资源。对一次文献信息进行加工、整理后的产品,即把大量的、分散的、无序的一次文献信息资源收集起来,按照一定的方法进行加工整理,使之系统化而形成的目录、索引和文摘等。因此,二次文献信息资源仅是对一次文献信息资源进行系统化和压缩,具有汇集性、检索性等特点。它的重要性在于提供了一次文献信息资源的线索,通过它可以节省查找一次文献信息资源的时间。

d. 三次文献信息资源。根据一定的目的和需求,对大量一次、二次文献信息资源进行筛选、分析、提炼、重组并再度出版的各种述评、手册、年鉴、百科全书等,具有参考性强、使用价值高等特点。

可以看出,从零次文献信息资源到一次、二次、三次信息资源,是一个从分散到集中,从无序到有序的对知识信息进行不同层次加工的过程。每一过程所包含的知识信息的质和量都不同,对人们查找信息所起的作用也各不相同。

(4)网络信息资源。网络信息资源是指通过计算机网络可以利用的各种信息资源的总和。具体来说是指所有以电子数据形式把文字、图像、声音、动画等多种形式的信息存储在光、磁等非纸介质的载体中,并通过网络通信、计算机终端等方式再现出来的信息资源。

信息资源由纸张上的文字变为磁性介质上的电磁信号或者光介质上的光信息,使信息的存储、传递、查询更加方便。网络空间存储的信息密度高、容量大,既可以在计算机内高速处理,又可以通过信息网络进行远距离传输。并且,网络信息的传递和反馈更加快速灵敏,具有动态性和实时性等特点。上传到网上的任何信息,瞬时就能传递到世界各地的每一个角落。同时,网络的共享性与开放性又使得每个人都可以在互联网上自由获取和存放信息,由于缺少质量控制和管理机制,这些信息没有经过严格的编辑和整理,参差不齐,形成了一个纷繁复杂的信息世界,给用户选择、利用信息带来了障碍。

第2章 信息检索基本知识

2.1 信息检索的概念和类型

2.1.1 信息检索的概念

信息检索(information retrieval)是指将信息按照一定的方式组织起来,并根据信息用户的需求查找出有关信息的过程。它的全称是“信息存储与检索”,这是广义的信息检索的概念。狭义的信息检索仅指该过程的后半部分,即从信息集合中找出所需信息的过程,相当于通常我们所说的信息查询。

信息的存储指对一定范围或特定专业的信息进行选择、提取、综合、特征描述等加工使其序化,即建立有序性的信息集合。检索是指借助于一定的设备与工具,采用一系列方法与策略从信息集合中查找出所需信息。存储是检索的基础,检索是存储的逆过程。传统的信息检索主要根据文献的内、外部特征,用手工方式实现,费时费力。现代以计算机为核心的信息检索技术,开辟了信息处理与信息检索的新时代。从单一处理文本信息发展到处理各种类型的超文本信息,这一过程不断拓展着信息检索的领域,丰富着信息检索的内容,大大提高了检索的效率。从本质上来说,信息检索就是指人们通过制订检索策略,利用一定的检索技巧和检索方法,从特定信息集合中迅速、准确地查找到自己所需要的信息,而不论它的类型和表现形式。

2.1.2 信息检索的类型

根据不同的标准,信息检索可划分为不同的类型。下面介绍两种目前比较普遍的划分方法。

(1)根据检索对象的不同,信息检索可分为文献检索、数据检索和事实检索。其中,文献检索是最基本、最主要的方式。

①文献检索(document retrieval)是以文献为检索对象的信息检索,即利用相应的方法和手段,在存储文献资源的检索工具或数据库中,查询用户在特定时间和条件下所需文献的过程。凡是查找某一主题、时代、地区、著者、文种的有关文献,以及回答这些文献的出处和收藏地点等,都属于文献检索的范畴。它为用户提供的是与用户需求相关的文献信息。例如,查找某一研究课题一定年限内的有关文献,或对一项发明进行查新,或从事新产品开发时查找有关最新研究动态等,均属于文献检索。文献检索是一种相关性检索,检索结果是文献线索,一般要查阅全文后才能决定取舍,与数据检索、事实检索相比,文献检索产生较早,检索系统也相对比较完备。

②数据检索(data retrieval)是指以数值或图表形式表示的数据为检索对象的信息检索,也称数值检索。例如查找某一数据、公式、图表、价格、化学物质的分子式、设备型号与参数等,都属于数据检索。数据检索是一种确定性检索,检索的结果通常是准确、可靠的数值或数据,可直接使用。比如查找世界上最高的山峰的准确高度、TNT(三硝基甲苯,一种烈性炸药)的结构式(见图 2-1)等。完成数据检索主要借助于各种手册、数值数据库、统计数据库等。

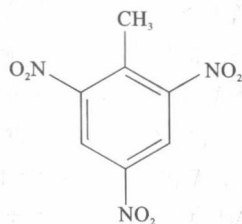


图 2-1 三硝基甲苯结构式

③事实检索(fact retrieval)是指对特定事件或事实的检索,包括事物的性质、定义、原理,以及发生的地点、时间、前因后果等。其检索对象既包括事实、概念、思想等非数值信息,也包括一些数据信息,但要针对查询要求,由检索系统进行分析、处理后再输出最终结果。事实检索是信息检索中最复杂的一种,要求检索系统必须有一定的逻辑推理能力和自然语言理解功能。目前,许多事实检索课题仍需要靠人工完成,但已有一些实验性的计算机事实检索系统。事实检索也属于一种确定性检索。

(2)根据检索方式划分,信息检索可分为手工检索和计算机检索。

①手工检索(hand retrieval)是指用人工来处理 and 查找所需信息的检索方式。它依靠检索者手翻、眼看、大脑判断进行,不需要借助复杂的检索设备。例如传统图书馆在实现数字化管理之前,主要靠目录柜等存储馆藏文献信息,读者要想查找自己所需的图书,必须通过手工检索来实现。手工检索的优点是直观、方便、灵活,可随时修改检索策略,查准率较高;缺点是检索速度慢,查全率低,不便于进行复杂概念课题的检索。

②计算机检索(computer retrieval)是利用计算机和一定通信设备查找所需信息的检索方式。它需要计算机、通信硬件设施和相应的应用软件。利用这种方式能对大量的信息进行存储,并可以根据用户要求从已存储的信息中迅速抽取特定信息。它的特点是速度快、效率高。不足之处是对软硬件和技术的要求较高,查准率较低。目前广泛使用的计算机检索系统包括光盘检索系统、联机检索系统和互联网上的搜索引擎。

2.2 信息检索工具

2.2.1 检索工具概述

为了从大量信息中及时获取特定需要的信息,必须借助于一定的检索工具。检索工具正是在人们解决庞大的信息和对信息的特定需求之间的矛盾中应运而生的,并且随着这种矛盾的逐步加深,检索工具的功能不断完善,应用领域不断拓展。

(1)检索工具的概念。

检索工具是指人们用来存储、报道和查找信息的工具;具体地说,就是汇集各种信息并按照特定方法编排,以供查考的工具或系统。

作为检索工具,它具有存储和检索两方面的基本功能。存储功能,指检索工具把汇集的有关信息按照其特征记录下来,使之成为一条条信息线索,并将它们序化,这就是所谓的信息存储过程。检索功能,指检索工具提供一定的检索入口,使人们能够按照一定的检索方