



21st CENTURY
实用规划教材

21世纪全国高等院校实用规划教材

现代信息检索

主编：樊爱国 薛德钧
主审：王金平

Xiandai Xinxi Jiansuo

Quanguo

Gaodeng

Shiyong Guihua Jiaocai



北京大学出版社
PEKING UNIVERSITY PRESS

G354
30

21 世纪全国高等院校实用规划教材

现代信息检索

主 编 樊爱国 薛德钧

副主编 杨桂荣 陈守鹏

左文革 林 红

编 委 (按姓氏笔画排序)

左文革 孙彩霞 陈守鹏 杨仁英

杨桂荣 张 勇 李剑飞 林 红

谢 康 樊爱国 潘素珠 薛德钧

主 审 王金平



北京大学出版社
PEKING UNIVERSITY PRESS

内 容 提 要

本书是 21 世纪全国高等院校实用规划教材。本着理论与实践相结合的原则,系统介绍了信息的理论基础、信息检索的原理、信息检索的途径和方法,还分别介绍了中文检索工具书、外文检索工具书的使用方法与检索技巧。为了适应新世纪社会信息化与现代信息技术迅猛发展的需要,着重讲述了中文数据库、外文数据库及 Internet 信息的检索方法及需要注意的事项。各章节理论讲述与检索实例的介绍相结合,既充分注重教材的基础性和实践性,又强调教材的科学性和先进性,使本教材做到“教师易教、学生乐学、实训好做、技能实用”。

本书既可作为全国高等院校各专业学生信息检索课程的教材,也可作为教学、科研人员继续教育的教材,还可作为广大信息用户与企业的科研人员、管理人员、工程技术人员从事文献信息管理应用的实用参考书。

图书在版编目(CIP)数据

现代信息检索/樊爱国,薛德钧主编. —北京:北京大学出版社, 2006.1

(21 世纪全国高等院校实用规划教材)

ISBN 7-301-10233-X

I. 现… II. ①樊…②薛… III. 情报检索—高等学校—教材 IV. G252.7

中国版本图书馆 CIP 数据核字(2005)第 140714 号

书 名: 现代信息检索

著作责任者: 樊爱国 薛德钧 主编

责任编辑: 李 虎

标准书号: ISBN 7-301-10233-X/G · 1815

出 版 者: 北京大学出版社

地 址: 北京市海淀区成府路 205 号 100871

网 址: <http://cbs.pku.edu.cn>, <http://www.pup6.com>

电 话: 邮购部 62752015 发行部 62750672 编辑部 62750667

电子信箱: pup_6@163.com

排 版 者: 北京东方人华北大彩印中心 电话: 62754190

印 刷 者: 河北滦县鑫华书刊印刷厂

发 行 者: 北京大学出版社

经 销 者: 新华书店

787 毫米×1092 毫米 16 开本 16 印张 370 千字

2006 年 1 月第 1 版 2006 年 1 月第 1 次印刷

定 价: 25.00 元

前 言

21 世纪,人类由工业社会全面进入信息社会。在这科学技术日新月异、知识信息爆炸的时代,终身学习越来越受到世界各国的高度重视。当代大学生肩负着社会主义现代化建设的重大历史使命,必须做到与时俱进、不断更新知识、改变原有的知识结构、努力提高自身的创新能力和信息素养,才能更好地迎接挑战。信息教育正是时代的要求,是当代大学生进行素质教育不可或缺的重要组成部分。大学生信息教育已经成为高等教育中越来越重要的一项任务。

《现代信息检索》是一门实践性很强的方法课,是全国高等院校信息检索课程系列教材之一。自 20 世纪 80 年代初国家教育部发出“关于在高等学校开设文献检索与利用课的意见”文件以来,许多院校根据文件精神相继开设了该课程。经过多年的教学实践证明,信息检索对培养在校学生的科技信息能力起到了重要的作用。然而,随着现代信息技术、网络通信技术的发展,以及国际互联网在全球的迅速普及,使人们搜寻和获取信息的方式和手段都发生了根本性的改变。这种变化不仅要求当代大学生具备应用现代信息技术、网络通信技术获取各种知识和信息的能力,而且对信息检索课的教学内容、教学手段和方法的改进与创新提出了更高的要求。基于此,为满足当代大学生进行信息素质教育的需要,我们结合信息检索课程教学革新的现状,在参考了大量相关文献的基础上,编写了本教材。

在编写过程中,我们着重强调本教材的通用性、普遍性、新颖性和实用性。做到:①在选材上涉及理、工、医、农等各个学科,体现了通用性特点;②对基础理论的表述深入详尽,体现了普遍性特点;③对检索工具检索原理和方法的介绍都以编写时的最新信息为准,突出了新颖性特点;④在内容上相对压缩了手检知识的比例,突出基本知识和方法应用的介绍,扩充和完善了机检知识的比例,特别是对搜索引擎、数据库和专业网站等方面的知识做了深入的介绍,强调了教材的实用性,读后使人有一种全新的感觉。

参加编写本教材的有中国农业大学、南昌大学、江西理工大学、南京中医药大学、江西中医学院和江西科技师范学院等高校中具有丰富教学经验的专家教授。本书共分 8 章,第 1 章由薛德钧、张勇编写;第 2 章由樊爱国编写;第 3 章由罗佳佳、杨仁英编写;第 4 章由杨桂荣、林红编写;第 5 章由李剑飞编写;第 6 章由陈守鹏编写;第 7 章由林红、潘素珠编写;第 8 章由左文革、孙彩霞、巢国正、陈士吉编写。全书最后由樊爱国、薛德钧统稿,王金平审定。

本教材在编写过程中参考了大量论著,对提高本教材的质量起到了重要作用,特向这些论著作者致谢。

由于编写人员水平有限,不足之处在所难免,恳请同行专家、师生和读者批评指正,以利再版时修定。

21 世纪全国高等院校实用规划教材

《现代信息检索》编委会

2005 年 10 月

目 录

第1章 绪论.....	1	第3章 中文检索工具书.....	36
1.1 信息及其相关概念.....	1	3.1 书刊检索工具.....	36
1.1.1 信息.....	1	3.1.1 书刊检索工具.....	36
1.1.2 知识.....	2	3.1.2 检索举例.....	41
1.1.3 情报.....	2	3.2 期刊论文检索工具.....	42
1.1.4 文献.....	3	3.2.1 期刊论文检索工具.....	42
1.1.5 信息、知识、情报、文献 的关系.....	6	3.2.2 检索举例.....	46
1.2 信息的特点与类型.....	7	3.3 特种文献检索工具.....	47
1.2.1 信息的特点.....	7	3.3.1 专利文献检索工具.....	47
1.2.2 信息的类型.....	8	3.3.2 标准文献检索工具.....	58
1.3 信息资源的特征与类型.....	9	3.3.3 会议文献检索工具.....	61
1.3.1 信息资源的特征.....	9	复习思考题.....	63
1.3.2 信息资源的类型.....	10	第4章 外文检索工具书.....	64
1.4 信息检索的意义和作用.....	11	4.1 美国《化学文摘》.....	64
复习思考题.....	13	4.1.1 概况.....	64
第2章 信息检索基础.....	14	4.1.2 CA文摘本的编排结构与 著录格式.....	64
2.1 工具书概述.....	14	4.1.3 CA索引系统和使用方法.....	68
2.1.1 参考工具书.....	14	4.1.4 CA指南.....	71
2.1.2 检索工具.....	21	4.1.5 CA的检索途径与检索程序.....	72
2.1.3 检索工具与参考工具书的 异同.....	25	4.1.6 CA的检索实例.....	73
2.2 检索语言和检索系统.....	25	4.2 美国《生物学文摘》.....	74
2.2.1 检索语言.....	25	4.2.1 概况.....	74
2.2.2 检索系统.....	29	4.2.2 BA的编排结构与著录格式.....	75
2.3 检索原理、方法、途径和步骤.....	30	4.2.3 BA的检索途径与检索程序.....	77
2.3.1 检索原理.....	30	4.2.4 BA的检索实例.....	78
2.3.2 检索方法.....	31	4.2.5 BA的相关报道.....	80
2.3.3 检索途径.....	32	4.3 美国《工程索引》.....	81
2.3.4 检索步骤.....	34	4.3.1 概况.....	81
复习思考题.....	35	4.3.2 EI的编排结构与著录格式.....	81
		4.3.3 EI主题词表.....	83

4.3.4	EI 的检索途径与检索程序	84	6.1.1	中国期刊全文数据库	121
4.3.5	EI 的检索实例	85	6.1.2	中国优秀博硕士学位论文全文数据库	128
4.4	其他外文检索工具	85	6.1.3	中国重要报纸全文数据库	128
4.4.1	美国《科学引文索引》	86	6.1.4	中国重要会议论文全文数据库	129
4.4.2	英国《科学文摘》	88	6.1.5	CAJ 全文浏览器	129
4.4.3	英国德温特公司专利检索工具	90	6.2	维普信息资源	133
4.4.4	《科技会议录索引》	93	6.2.1	《中文科技期刊数据库》(文摘版、全文版)	133
	复习思考题	94	6.2.2	《中文科技期刊数据库》(引文版)	141
第 5 章	计算机信息检索基础	95	6.2.3	《外文科技期刊数据库》(文摘版)	141
5.1	概论	95	6.2.4	《中国科技经济新闻数据库》	142
5.1.1	计算机信息检索的概念	95	6.2.5	维普行业信息资源系统	142
5.1.2	计算机信息检索发展概况	95	6.3	万方数据资源系统	142
5.1.3	计算机信息检索特点	97	6.3.1	概况	143
5.1.4	计算机信息检索意义	98	6.3.2	检索途径和方法	143
5.2	计算机信息检索系统	99	6.3.3	检索结果的输出与处理	148
5.2.1	计算机信息检索系统的物理构成	99	6.4	其他中文数据库	150
5.2.2	数据库的类型	100	6.4.1	《人大复印报刊资料》数据库	150
5.2.3	数据库的结构	102	6.4.2	《全国报刊索引数据库》	153
5.3	计算机信息检索服务的类型	104		复习思考题	156
5.3.1	光盘信息检索	104	第 7 章	外文数据库	158
5.3.2	联机信息检索	106	7.1	SpringerLink 全文数据库	158
5.3.3	网络信息检索	108	7.1.1	概况	158
5.4	计算机信息检索基本技术	110	7.1.2	数据库的登录和主页介绍	158
5.4.1	布尔逻辑检索	111	7.1.3	检索途径和方法	159
5.4.2	截词检索	112	7.1.4	检索结果的输出与处理	163
5.4.3	字段检索	113	7.2	OVID 数据库系统	164
5.4.4	全文检索	114	7.2.1	概况	164
5.4.5	其他检索技术	115	7.2.2	检索途径和方法	165
5.5	计算机信息检索效果评价	116	7.2.3	检索结果的输出与处理	169
5.5.1	评价目的和范围	116	7.3	EBSCOhost 数据库	171
5.5.2	评价标准	116	7.3.1	概况	171
5.5.3	影响检索效果的因素	118	7.3.2	检索途径和方法	172
	复习思考题	119			
第 6 章	中文数据库	120			
6.1	中国知识基础设施工程	120			

7.3.3 检索结果的输出与处理.....	177	8.2.2 中文搜索引擎.....	198
7.4 其他外文数据库	178	8.2.3 英文搜索引擎.....	205
7.4.1 OCLC FirstSearch 检索系统...178		8.3 网上信息检索.....	208
7.4.2 Elsevier SDOS 全文 电子期刊	179	8.3.1 概况.....	208
7.4.3 ProQuest 检索系统.....	179	8.3.2 书目信息检索.....	210
7.4.4 BIOSIS Previews 数据库	180	8.3.3 电子期刊检索.....	217
7.4.5 CA on CD 数据库.....	181	8.3.4 电子图书检索.....	220
7.4.6 EI Compendex Web 数据库	181	8.3.5 电子报纸检索.....	225
复习思考题	182	8.3.6 专利信息检索.....	228
		8.3.7 其他信息检索.....	230
		复习思考题.....	230
第 8 章 Internet 信息检索	183	附录	232
8.1 Internet 概述	183	附录 I 《中国图书馆分类法》(第四版) 简表	232
8.1.1 Internet 基本概念	184	附录 II 《中国标准文献分类法》简表	236
8.1.2 Internet 基本原理	186	附录 III 常用网址.....	240
8.1.3 Internet 提供的服务形式	188	参考文献	246
8.1.4 计算机网络安全.....	194		
8.2 搜索引擎及其使用技巧.....	196		
8.2.1 搜索引擎类型与特点.....	196		

第1章 绪 论

在文字发明以前,人类文明的传承只能通过一代代人的口头相传,当时,人们所获得的信息和知识的数量是有限的。随着人类利用信息范围的不断扩大和对自然认识的不断深入,不同的人群发明了不同的文字来记录各自的生活经验。直到纸张和印刷术发明及广泛应用之后,世界上的文献数量急剧增加,人们之间的信息交换和信息利用的能力也逐步增强,各种信息检索技术应运而生。到20世纪前期,各种信息检索技术逐步成熟,并逐步形成了一门专门的学科,即信息检索,来研究各种已有信息检索技术的应用和开发新的信息技术。

21世纪,随着社会的进步、科学的发展、新技术的应用,尤其是计算机网络化、全球化进程加快,社会信息量呈指数级增长,知识和信息已经成为推动科技进步和社会发展的决定性因素。人们的信息意识、信息观念在不断更新,经济技术和科学知识的创新能力在不断增强。海量的信息资源,多样的信息类型和表现形式以及复杂的时空关联关系,为信息的开发与利用提供了便利的条件,也为信息的发布与共享提供了外部环境。同时,信息产生和传递的随机性、信息时空关系和检索系统状态的不确定性,以及信息资源的庞杂同特定需求之间的矛盾,也给人们搜集与利用信息增加了困难和不便。为了更准确、更快捷地从信息的海洋中找到所需的信息,我们首先应了解信息及其相关概念。

1.1 信息及其相关概念

时至今日,人类不仅需要对生活于其中的世界进行不断深入的了解,而且不同的人群之间需要协作与沟通,也需要加强相互之间的通信与交流,信息、知识、情报和文献越来越重要地影响着人们的生活,因此,准确全面地了解它们的含义非常有必要。

1.1.1 信息

信息(Information)一词最早见于我国南唐诗人李中《暮春怀故人》中“梦断美人沉信息,目穿长路倚楼台”。诗中信息意指消息。随着社会的不断发展,人们对信息的认识不断深化,许多知名学者都试图给出信息的准确含义。如控制论的创始人维纳(Norbert Wiener)认为:信息是我们对外部世界进行调节并且使这种调节为外部世界所了解的过程中同外部世界进行交流的内容的名称。而信息论的奠基者、美国数学家申农(C. E. Shannon)则从研究通信理论出发,在其《在噪声中的通信》一书中将信息定义为:信息是不确定性的减少或消除。

我国学者杨文祥从不同学科角度对信息概念进行了概括:

(1) 从信息论、控制论学科角度 其代表就是申农所给出的信息定义。他在研究广义通信系统理论时还指出:信息量就是信息源的不定度。

(2) 从哲学角度 当代哲学界关于信息的概括性描述已取得了一定程度的共识:信息是系统有序程度的度量。

(3) 从信息管理学角度 信息管理专家霍顿(F. W. Horton)给信息下的定义是:信息是按照最终用户决策的需要,经过处理和格式化的数据。处理可以是自动化的或手工的,由数据转化为信息是由信息处理者自己完成的。

(4) 国家标准对信息的定义 我国《情报与文献工作词汇基本术语》(GB48944-1985)中关于信息的定义是:物质存在的一种方式,一般指数据、消息中所包含的意义,可以使消息中所描述的事件的不定性减少。

从各种对信息定义的阐述来看,信息不是事件或物体本身,信息是事件或物体的某种属性的反映,是客观事物某些特性的表象或线索。因此,我们认为,信息是世界上一切事物的运动状态、特征及其反映。

1.1.2 知识

知识(Knowledge)是人类在改造世界的过程中所获得的认识和经验的总和。人们把在利用自然、改造自然的社会实践中所搜集和利用的信息加以概括与总结,发挥自身的主观能动性,对所掌握的信息加以系统化和完备化,就生成了具有一定模式的知识。因此,知识是人类大脑通过思维重新组合的信息,是系统化完备化的信息。例如,彗星俗称扫帚星,其形状古怪,变化多端,行踪诡秘,它的出现在古代常被视为不祥之兆。这是因为绝大部分彗星的运行周期都在成百年以上,前人虽然曾多次看到彗星并加以记载,但他们所能认识到的只是自然界中客观事物的表现,记载下的只是客观世界中的自然信息。由于当时人们的这种感性认识尚不能形成知识,所以他们无知地把彗星视作不祥之物就不足为怪了。直到哈雷运用牛顿力学的规律,成功地预言哈雷彗星的运行周期时,人类对彗星的认识才从感性认识上升到理性认识。这种理性认识就是对客观事物全面的、系统的、规律性的认识,这样的认识便是知识。因此,知识才是破除迷信的最有力武器。

按照知识的职能可分为隐性知识和显性知识两类。国际经济合作与发展组织(OECD)在1996年发表的《以知识为基础的经济》报告中,按知识的表现形态将其分为四大类:know what,即事实知识;know why,即原理知识;know how,即技能知识;know who,即人力知识。后来又有人补充了两类:know where即定位知识和know when即定时知识。知识具有不可替代性、不可分割性、不可相加性、不可逆转型、可共享性、无限增值性、非磨损性等特征。

知识和信息如果仅仅是拥有,未必能取得效益,只有将信息转化为知识,再将知识转化为智慧,即将信息和知识应用到创新中才能产生强大的生产力。这是一种动态的过程、创新的过程。创新的结果是将经过信息到知识共享后得到的隐性知识上升为显性知识并付诸实践,成为智慧;反过来,智慧又常会转化为新的知识,新的知识又变成新的信息,通过一定的手段和社会传递过程,借助于某些载体传递给特定的使用者。

1.1.3 情报

古时候,人们对于情报(Information)的理解是“战时关于敌情之报告”。随着时间的变化,现在所说的情报已大大突破这个狭隘的限度。我国《辞海》将情报定义为:一是以侦

察手段或其他方法获得有关敌人的军事、政治、经济等各方面的情况,及对这些情况进行分析研究的成果,是军事行动的重要依据之一;二是泛指一切最新的情况报道。作为图书情报学意义的情报一词是舶来品,来自日本。关于情报的定义,国内外学术界至今仍没有定论,但大家的基本共识为:情报是知识经传递并起作用的部分,或者说情报就是运用一定的形式传递给特定用户并产生效用的知识。由此,可以看出情报具有三个基本属性:知识性、传递性和效用性。

1. 情报的知识性

情报来源于知识,情报具有特定的知识内容,且是以运动形式存在的,是在特定条件下对特定对象有用的那一部分知识。

2. 情报的传递性

无论多么重要的知识,人们不知道其存在就不能运用其来解决特定的需求,因而也不能称其为情报。知识要成为情报,必须经过运动,传递给特定的对象并满足其特定的需求。所以说情报是激活了的知识。

3. 情报的效用性

人们创造情报、传递情报的目的,在于利用情报。情报的效用性表现为启迪思想、开阔视野、增进知识、改善人们的原有知识结构、提高人们认识世界和改造世界的能力、充分发挥情报的经济效益和社会效益。

在知识创新、竞争激烈的现代社会,人们在生产实践、科技开发、商业经营等社会活动中,都离不开情报。情报已经广泛渗透到社会各领域:政治情报、军事情报、经济情报、科技情报,构成了一个国家的情报体系,成为发展经济、巩固国防、增强综合国力、提高国防竞争力的重要条件。

1.1.4 文献

1. 定义

世界上最早的文献(literature)产生于距今大约 10 000~40 000 年前的旧石器时代。雕刻在洞壁上的石刻图像、古代西亚的楔形文字文献、古埃及的纸草文献、古希腊克里特线形文字文献和中国商代的甲骨文文献等,都是世界上最古老的文献。在我国,“文献”一词最早见于《论语·八佾》。子曰:“夏礼吾能言之,杞不足徵也;殷礼吾能言之,宋不足徵也;文献不足故也。”宋代朱熹在《四书章句集注》中注释为:“文,典籍也;献,贤也。”这里,“文”指典章制度的文字资料,“献”指见多识广、熟悉掌故的人。元代马端临在《文献通考·总序》中对“文献”的诠释是:“凡经、史、会要、百家传记藏书,信而有证者,谓之文;凡臣僚之奏疏,诸儒之评论,名流之燕谈,稗官之记录,一语一言,可以订典故之得失,证史传之是非者,谓之献。”这里所说的“文献”指文字资料和言论资料。可见,随着时代的变迁,文献的概念也发生了变化,原来所含“贤人”之意逐渐消失了。

到了现代,随着科学的发展,文献的外延扩大了。国际标准化组织颁布的《文献情报术语国际标准》(ISO/DIS 5217)对文献的解释是“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。”

我国 1983 年颁布的国家标准《文献著录总则》(GB3792.1-1983)中,对文献的定义为“文献是记录有知识的一切载体。”由此可以看出,文献是由知识内容、信息符号和载体材料三个不可分割的基本要素构成。知识是文献的实质内容,符号、文字、声音是人体感觉信息的媒介;载体是文献的外在形式;同时,附载有知识内容的信息符号必须通过一定的记录手段存附在一定载体上才能形成文献。

2. 类型

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。

1) 按照对文献的加工程度分

(1) 一次文献 作者以本人进行的科技活动或生产活动中所取得的成果为基本素材而创作的原始文献。如期刊论文、会议论文、科技报告、学位论文等。这类文献记载的是具有原创性质的知识信息,具有较高的情报价值。但其数量庞大,发表分散、无序。一次文献是检索的最终目的,文献检索就是把最适用的具体的一次文献查找出来。

(2) 二次文献 是指文献信息工作人员对无序的一次文献进行加工整理后形成的文献。如各种书目、索引、题录等。二次文献是在对一次文献的处理的基础上产生的,编制目的是为一次文献的快速和准确查找提供服务。因此,二次文献并不以反映一次文献的学术内容为重点,而是通过对一次文献的外部和内容特征的标示,从而系统地反映一次文献的检索线索。二次文献是查找一次文献的重要工具,学习文献检索主要要掌握二次文献的利用方法。

(3) 三次文献 是指在合理利用二次文献的基础上,选用一次文献的内容进行分析综合而编写出的文献。主要包括两种类型:一类旨在对某些具体研究课题的当前状况、发展趋势进行分析和评述,如综述、专题评论等;另一类是将发展较为成熟的知识系统化,以便人们学习、查找和利用,如教科书、词典、年鉴、百科全书、指南等。三次文献是以一次文献为研究对象,二次文献为研究工具,并通过各种情报研究方法而产生的更高层次的情报研究型文献,无论是知识成熟程度、概念表达的准确性,还是其见解或结论的客观性,都达到较高水平。三次文献一般可以反映出该学科的概貌,具有很强的资料性和应用指导作用。三次文献一般附有大量的参考文献。

(4) 零次文献 又称为准文献。它是指未来得及记录下来、尚未发表或不适合公开和大范围交流的比较原始的素材、底稿、手稿、工作文稿、原始统计数字、工程图纸、实验记录以及各种口头交流的经验、意见和各种操作示范性方面的知识等。零次文献一般在较小的范围内交流、使用、参考,其内容一般比较新颖,具有较高的价值,难以获取。

2) 按照文献的出版形式分

(1) 图书 现代印刷型出版物中最普通的一种文献,一般又可分为科技图书、文艺图书、科普读物、教科书、工具书等。图书的内容比较系统、成熟,是生产技术、科技成果的概括和社会生活的总结。

(2) 期刊 又称杂志。它是一种定期或不定期出版的连续出版物,有固定的期刊名、刊号和出版形式。期刊内容新颖、信息量大、报道迅速,能够反映当前的科技水平和社会发展,是主要情报信息源之一。

(3) 专利文献 专利是由政府有关部门(如国家专利局)根据发明人的申请,认为其发明

符合法律规定的条件而授予发明创造者在一定时期内对其发明独有的制造、使用和销售的权利。专利文献主要指发明人呈交的说明该项发明的目的、技术梗概和专利权限等的专利申请书和专利说明书以及专利主管部门审查批准后予以公布的专利公报等。专利文献一般附有图表、数据和较为详细的文字说明。专利的性质是发明与创新,反映了当时科研、开发的最新成就,因此专利文献是重要的情报信息来源。

(4) 会议文献 指在专业学术会议上宣读或书面交流的论文、学术报告、会议纪要等文献。它反映某学科或专业的最新研究成果和发展趋势,是了解学术动态、研究进展的重要信息源。

(5) 学位论文 指高等院校、科研机构的毕业生为取得学位在导师指导下完成的学术性研究论文。学位论文是带有一定创造性的一次文献,对科研有一定参考价值,常成为大学图书馆的特藏。

(6) 标准文献 又称标准资料,是指有产品和工程的质量、规格、生产过程、检验方法等的技术文件。标准具有一定的法律约束力,如任务说明书、协议书、技术指标、实验计划、病案资料等。标准文献文字简练、词义准确,一项标准仅针对一个技术问题,提供的数据和技术内容具体而严密,有很强的针对性和指导性。

(7) 其他 如技术档案、产品资料、报纸、政府出版物等,具有报道快、涉及面广、内容新等特点,是人们获取信息的重要途径。

3) 按照文献的载体形式分

(1) 手写型文献 主要指印刷术发明之前的古代文献和当今没有正式付印的手写记录。它以刻画或手写为手段,将知识内容记录在各种自然材料或纸张等载体上,如古代的甲骨文、金石文、简策、帛书以及现代的笔记、手稿、书信等。

(2) 印刷型文献 是以印刷技术(如铅印、胶印、油印等)为手段,以纸张为记录信息内容的载体而产生的一种文献形式。如书、刊及卡片等。印刷型文献以成本低廉、携带方便、可直接阅读、比较符合人们的传统阅读习惯等优点在现阶段文献类型中仍占主导地位。但也具有存储密度低、体积大、保存期限短等缺点。

(3) 缩微型文献 采用照相技术,将文献按一定比例缩小存储在感光材料上形成的一种文献。如缩微胶卷、缩微平片等。其特点是体积小、分量轻、容量大、保存期长、传递方便,但阅读不方便,必须借助阅读机才能阅读。

(4) 声像型文献 又称视听型文献。它是把知识以声音或图像等形式记录在唱片、录音带、录像带、影片、幻灯片等视听载体上,再通过录放装置,以声图并茂的形式展现的一种文献。其特点是传递知识直观、形象、生动、逼真,宜于记载难以用文字表达和描绘的形象资料和声频资料。

(5) 电子型文献 又称机读型文献、数字化文献。它是通过编码程序设计把文献变成数字语言和机器语言输入计算机并存储在磁介质或激光介质上,阅读时再利用计算机输出的一种文献。如电子图书、电子期刊、电子新闻等。电子型文献不仅有很高的信息存储密度,而且还有高速的信息存取速度,并具有电子加工、出版和传递功能,宜于实现资源共享。

随着网络出版物的大量盛行,近来有人提出将完全依靠计算机网络传递的电子文献从电子型文献中划分出来,作为独立的网络型文献,如联机数据库、网络上的软件、游戏等。其划分的主要依据为文献传递和复制的方式不同。电子型文献采用拷贝等方式以实物形式

传递, 而网络型文献则主要采用下载或套录(Download)的方式依靠网络传递。我们认为, 网络型文献信息资源从其本质上来说, 也属于机读型文献信息资源, 故不将其作为一种文献形式介绍。

1.1.5 信息、知识、情报、文献的关系

信息、知识、情报、文献之间既有区别又有联系。信息广泛存在于自然界和人类社会, 其涵盖面最广; 信息被人类所感知并被提炼加工即成为知识; 情报是激活了的知识, 是特定用户为一定的效用目的而获取的知识; 文献则是记录知识的载体。它们之间的逻辑关系和范畴关系分别如图 1.1 和图 1.2 所示。

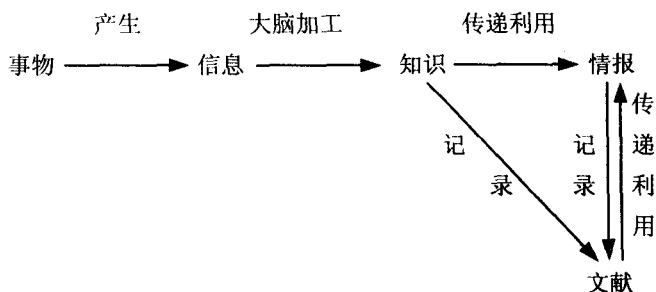


图 1.1 信息、知识、情报、文献的逻辑关系

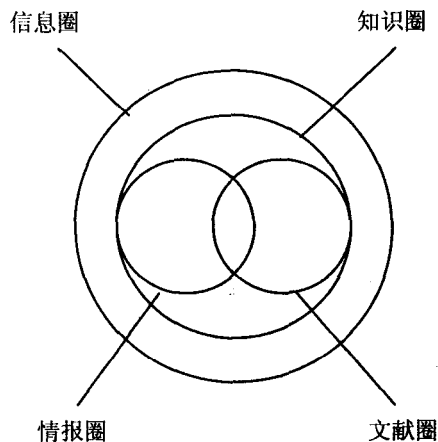


图 1.2 信息、知识、情报、文献的范畴关系

一方面, 人们将信息、知识、情报经过选择、分析、综合、记录和存储在一定载体上, 加工成文献, 这是文献的创作过程; 另一方面, 人们通过不同类型的文献认识和掌握纷繁复杂的信息、知识和情报, 用于认识世界和改造世界, 这就是文献的利用过程。

目前, 国内学者对信息和情报的关系持有不同见解: 由于信息和情报都对应于英文 information 一词, 国外并不区分二者, 因此有人主张用信息代替情报, 认为有利于国际交流; 但也有人认为这种替代并不符合客观实际, 因为二者是不同的两种概念, 所以不能互相替代。

1.2 信息的特点与类型

信息定义的多样化,主要有3个方面的原因:(1)信息本身的复杂性;(2)对信息的研究尚不够完善成熟;(3)实际需要的不同。由于前两个原因,人们出自不同的研究目的,从不同的角度出发,对信息做出不同的理解和分析。为了更好地认识信息,我们将对各种对信息的特点和类型的阐述进行整合。

1.2.1 信息的特点

由于目前对信息本质的认识仍存在分歧,因而对信息特点的描述更是说法众多。概括起来主要有以下几点:

1. 普遍性

世界是物质的,物质是运动的,运动的物质都产生和带有信息。物质的普遍性以及物质运动的规律性决定了信息的普遍存在性。

2. 客观性

信息是客观存在的,信息的客观性是以物质的客观存在为前提的,不受人的主观认识的限制。信息可以被感知、被处理和存储、被传递和利用。

3. 无限性

在人类生存和社会活动的一切领域,随时都在交流、产生和处理信息。物质的边界是无限的,人们对物质世界的认识也是无限的,人的创造力也是无限的,随着时间的推移,人类所掌握和涉及的信息内容和范围在扩大,信息量也在无止境地增加。

4. 传递性

我们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。传递性是信息的基本要素和明显特征。信息只有借助于一定的载体(媒介)经过传递才能为人们所感知和利用。

5. 可扩散性

信息的传递性决定了信息的可扩散性。可扩散性指信息通过各种渠道、媒介进行传播。信息网络的覆盖面越来越大,信息的扩散是无垠的。

6. 共享性

信息源于事物,但又不是事物本身,是人们用来认识事物的媒介。信息具有共享性是区别信息不同于物质和能量的最主要特征。即同一内容的信息在同一时间、同一地域可以被两个或两个以上的不同受众所分享,其分享的信息量不会因受众数量的多少而受到影响,原有的信息量也不会因之改变。

7. 有用性

现在,信息已成为人类社会的重要资源,人们利用信息来认识和改造客观世界、创造

价值。如企业获得了足够信息后,可以做出正确的决策,获得更多的投资回报。证券市场更是如此。

8. 增值性

信息可经过无限多次的传递使用,在传递使用的过程中,其原本信息量不仅不会因之损失或减少,而且往往经过传递使用可以产生新的信息,即增值。

1.2.2 信息的类型

信息类型比较复杂,按照不同的标准,可对信息进行不同的分类。

1. 按信息源的不同划分

(1) 自然信息 即由自然界客观存在的固有事物所发出的信息。形形色色的天气变化、地壳运动、天体演变、自然奇观……自然界每时每刻都在散发着大量的信息,只是由于条件的限制,我们对于这类信息的认识还远远不够。

(2) 机电信息 即由机电设备所发出的信息。如电子仪器的脉冲信号、无线电波等传递的信息等。现在这类信息在我们的生活、科研等方面已成为不可或缺的因素。

(3) 生物信息 即由生物体发出的信息。如鸟语花香、体温升降等。有关试验研究表明,植物之间存在着信息交换现象,植物能够感知并传递信息。动物之间更有着特定的信息联系方式,各种动物都有自己交换信息的“语言”。而传递信息的作用则是生命进化的重要原因,没有信息,就没有丰富多彩的生物界,更不会出现人类社会。

(4) 社会信息 社会信息是指人类活动所产生及伴随的各类信息。其范围相当广泛,例如人口信息、法律与政治信息、教育状况的信息等。这些信息的收集、管理、加工与利用是各级各类政府部门、事业单位的主要任务。各国政府对于本国社会情况编制的年鉴、报告就是典型的例子。

按照活动领域,社会信息又可分为科技信息、经济信息、政治信息、军事信息、文化信息等。

- ◆ 科技信息:人类在进行科学研究中积累了大量的信息,包括各种理论、学说、发明、专利等大量的资料数据。这些信息既是几千年来人类研究客观世界的收获,又是进一步创造信息、改造世界的出发点与基础。现代的实验与观测手段,在收集与处理这些信息方面为人类提供了更多的帮助。无论是大型的加速器还是哈勃天文望远镜,从其功能来说都是为了人类能获取更多的信息。
- ◆ 经济信息:经济活动是人类最基本的社会活动。在生产、消费、流通、分配的经济活动全过程中产生、收集、处理及利用的信息总和即为经济信息。如果说在原始社会的物物交换中,信息还只是隐含在各种商品之中,那么随着经济活动的发展,信息的作用就越来越明显,信息越来越脱离了具体的载体,成为抽象的一般等价物。首先是货币,然后是各种有价证券,最后到现代的电子交换方式。经济信息是人类所处理的信息中最基础也是最多的一类信息。
- ◆ 文化信息:在人类社会基本的日常活动之中,形成了丰富多彩的文化生活,积累了众多的文化遗产,包括文字、诗歌、音乐、绘画、雕塑等大批信息。在传统意义上,图书馆是这些信息的集中地,而学校则是传播这些信息的主要场所;在现代技术条

件下,报纸、广播、电视、电影都在这面对社会产生了巨大的影响。

2. 按主体认识的层次划分

从主体对信息的认识层次上看,由于主体有感受力,能够感知事物运动状态及其变化方式的外在形式,由此获得的信息称为语法信息;由于主体有理解力,能够领会事物运动状态及其变化方式的逻辑含义,由此获得的信息称为语义信息;由于主体具有明确的目的性,能够判断事物运动状态及其变化方式的效用,因此获得的信息称为语用信息。语法信息、语义信息和语用信息三位一体的综合,即构成了认识论层次上的全部信息。

(1) 语法信息 语法信息是信息认识过程中的第一个层次。它只反映事物存在方式和运动状态,而不考虑信息的内涵。换言之,语法信息只是客观事物形式上的单纯描述,只表现事物的现象而不深入揭示事物发展变化的内涵及其意义。

(2) 语义信息 语义信息是信息认识过程中的第二个层次。它是指认识主体所感知或所表述的事物的存在方式和运动状态的逻辑含义。换言之,语义信息不仅反映事物运动状态的变化,而且还要揭示事物运动变化的意义。

(3) 语用信息 语用信息是信息认识过程中的第三个层次。它是指认识主体所感知或所表述的事物的存在方式和运动状态相对于某种目的所具有的效用。换言之,语用信息就是指信源^①发出的信息被信宿^②接收后将产生的效果和作用。

此外,根据信息的时态可将信息分为静态信息、历史信息和预测信息;依据信息的记录与否可将信息分为记录信息和无记录信息;按照信息对管理作用的大小可将信息分为战略信息和运行控制信息;按照信息的流动方向可将信息分为反馈信息和前馈信息等。无论按照何种标准划分,都是为了更清楚地认识信息和更准确地利用信息。没有物质,就没有我们生活的世界;没有能量,世界就将消亡;没有信息,物质和能量只能形成一个混浊、杂乱的空间。物质向人类提供材料,能量向人类提供动力,信息向人类提供的则是知识和智慧。信息资源与物质资源、能量资源一起,共同构成现代社会资源的三大支柱。

1.3 信息资源的特征与类型

目前,学术界对信息资源的定义有两种不同的观念:一是狭义论,认为信息资源就是指文献资源或数据资源,或指各种媒介和形式的信息集合。二是广义论,认为信息资源是信息活动中各种要素的总称,既包含了信息本身,也包括与信息活动相关的人员、设备、技术和资金等各种资源。无论从哪种角度,信息只有在实施管理后才具有成为资源的条件,所以说,信息资源是经过人类开发与组织的信息的集合。

1.3.1 信息资源的特征

信息资源主要是再生信息和一部分感知信息,因而,信息资源除“无限性”之外,具有信息的所有特点。相对于其他非资源型信息,信息资源具有四个明显的特征:

① 信源:信息发出者。

② 信宿:信息传递过程中的接收者,即接收信息的人或物。

1. 智能性

信息资源是经过人类开发和组织的信息,是人类脑力劳动或者说认知过程的产物。人类社会在特定时期内的整体智能水平决定着特定时期社会信息资源的量与质,个人的智能水平和信息意识也决定了个人所掌握的信息资源的量与质。

2. 有限性

经过人类开发与组织的信息只是无穷无尽的信息海洋中极有限的一部分,因此,相对于信息的无限性来说,信息资源是有限的;相对于人类永不满足的信息需求来说,信息资源也是有限的。信息资源的有限性特征要求人类必须从全局出发,合理布局 and 共同利用信息资源,最大限度地实现资源共享,从而促进人类与社会的发展。

3. 不均衡性

由于人们的认识能力、知识储备和所处的信息环境等多方面的条件不尽相同,他们所掌握的信息资源的多寡也不均;同时,由于社会发展程度不同,对信息的开发和组织的程度不同,地球上不同区域信息资源的分布也不均衡。通常所谓的信息领域的“马太效应”就是与这种不均衡性有关的现象。

4. 整体性

信息资源作为一个整体,是对一个国家、一个区域或一个组织的政治、经济、文化、技术等全面反映。信息资源的每一个要素只能反映某一方面的内容,割裂它们之间的联系则无异于盲人摸象。信息资源的整体性特征要求对所有的信息资源和信息资源管理机构实行统一的管理,而避免人为的分割所造成的资源建设的重复和浪费。

信息资源从本质上说是一种附加了人类劳动的信息。进一步分析,信息资源是由信息、人、符号、载体四种最基本的元素构成的。其中,信息是信息资源的源泉;人作为认知的主体是信息资源的生产者和利用者;符号是人生产和利用信息资源的媒介和手段;载体则是储存和利用信息资源的物质工具。换句话说,信息资源是人类通过一系列的认知和创造过程之后以符号形式储存在一定载体上可供利用的全部信息。

1.3.2 信息资源的类型

信息资源的类型可以根据多种标准来划分,通常按照以对其开发的程度为依据来划分,将信息资源分为潜在信息资源和现实信息资源两大类。

潜在信息资源是指个人在认识和创造过程中储存于大脑中的信息资源。它们虽然能为个人所利用,但一方面易于随记忆忘却过程而消失,另一方面又无法为他人所直接使用,因此是一种有限再生资源。现实信息资源是指潜在信息资源经个人表述之后能够为他人直接使用的信息资源。现实信息资源最主要的特征是具有社会性,通过特定的符号表述和传递,可以在特定的社会条件下广泛地连续往复地为人类所利用,因此是一种可无限再生的资源。

现实信息资源以表述方式的不同可分为口语信息资源、体语信息资源、实物信息资源和文献信息资源。口语信息资源是人类以口头语言所表述出来而未被记录下来的信息资源,