

信息检索与查询

XINXI JIANSUO YU CHAXUN

侯 戎 非 主编

合肥工业大学出版社



内容简介

信息检索与查询

主 编 侯戎非

副主编 吴显中

顾 问 费文升

编 者 侯戎非 廖 捷

吴显中 张先锋

主 编

合肥工业大学出版社

合肥市中环路193

230009

总编室:0551-2901301

发行部:0551-2901301

印 张 11.75

字 数 382千字

安徽新华书店

全国新华书店

主 编

合肥工业大学出版社

合肥市中环路

230009

总编室:0551-2903038

发行部:0551-2903038

www.hjpress.com.cn

press@hjpress.com.cn

ISBN 950-7-7481-1 定价:18.00元

如有任何质量问题,请与出版社发行部联系。

内容简介

随着计算机技术的迅速发展,信息正逐步从传统的纸制载体向电子载体发展,信息检索技术也将不再局限于传统的手工检索方式,而是以计算机技术为核心向着多元化方向发展。本书从符合读者实际使用角度出发,在信息检索基本知识体系阐述完整的基础上,着重对网络信息资源检索进行介绍。

全书共分信息与信息检索基本知识、传统信息检索技术、网络信息检索技术三大部分共七个章节。适用于高职高专类院校各专业学生,同时对广大信息检索知识爱好者也有一定的参考价值。

图书在版编目(CIP)数据

信息检索与查询/侯戌非主编. —合肥:合肥工业大学出版社,2009.5

ISBN 978-7-81093-949-2

I. 信… II. 侯… III. 情报检索—高等学校:技术学校—教材 IV. G252.7

中国版本图书馆 CIP 数据核字(2009)第 060529 号

信息检索与查询

侯戌非 主编

责任编辑 权 怡

出版	合肥工业大学出版社	版次	2009年5月第1版
地址	合肥市屯溪路193号	印次	2009年5月第1次印刷
邮编	230009	开本	787毫米×1092毫米 1/16
电话	总编室:0551—2903038 发行部:0551—2903198	印张	11.75
网址	www.hfutpress.com.cn	字数	282千字
E-mail	press@hfutpress.com.cn	印刷	安徽江淮印务有限责任公司
		发行	全国新华书店

ISBN 978-7-81093-949-2 定价:19.00元

如果有影响阅读的印装质量问题,请与出版社发行部联系调换。

序

进入信息社会时代,各种信息日益增多,单位与单位之间、人与人之间的交往和联系愈加频繁、密切,人们需要通过联系与沟通,以形成一个相互协调、有序运行的有机整体,共同促进人类与社会的进步与发展。

而随着现代电子技术、现代信息技术和现代网络技术的广泛应用,给人们的工作、生活带来了新的变化、新的活力和新的发展。人们需要充分利用现代科学技术手段来获取信息、筛选信息、存储信息和传递信息,以便更好地进行交流和沟通。

为适应现代社会发展的需要,人们迫切需要掌握信息检索与查询方面的知识与技能。基于此,本书作者以实际需要和实际应用为出发点,编写了本书。

本书较为系统地介绍了信息及信息检索的基本知识,包括信息资源、信息检索的工具和语言、信息检索的方法和步骤;按照现代信息检索技术的新要求,介绍了最新的网络信息资源检索与查询的手段、途径及其方法和步骤,从而使本书更具时代性、应用性及应用的广泛性特点。

本书编写的初衷,仅为文秘专业课程设置的需要而撰写的,但在实际教学工作中,高职院校不仅需要加强学生人文素质的教育与培养,而且还要加强学生科学技术素养方面的教育与培养,为此,我们深切地感到,该书的出版发行,不仅要满足文秘专业课程开设的需求,而且还要满足其他各专业开设该课程的需求,并成为其他专业的专业基础课或专业选修课,以便学生在今后的实际运用中,能更规范、更方便、更快捷地检索与查询所需要的信息资料,变知识为能力。同时,该书的出版发行,也适应并满足了非在校学习的相关人员工作与生活中对该方面知识与技能的需要和需求。

担任本书编写任务的老师们都已从教多年,先后承担过文秘、汽车技术服务与营销、护理及医药营销等专业该课程的教学任务,在教学信息反馈中,学生对开设该课程及该课程授课教师的教学水平及教学能力给予了充分肯定。尽管如此,在组织教学及撰写本书过程中,不可避免、或多或少地存在着一些问题与不足,在此,谨请各界人士给予必要的严厉批评与指正,并请各位专家给予更多的关注与支持。谢谢!

费文升

二〇〇九年四月

目 录

第一章 信息的基本知识	(1)
第一节 信息概述	(1)
第二节 文献信息资源	(6)
第三节 计算机信息资源	(13)
第二章 信息检索的基本知识	(19)
第一节 信息检索概述	(19)
第二节 信息检索工具	(22)
第三节 信息检索语言	(26)
第四节 信息检索的方法、途径和步骤	(33)
第五节 信息检索效果的评定	(37)
第三章 计算机信息检索	(40)
第一节 计算机信息检索概述	(40)
第二节 计算机信息检索技术	(45)
第四章 网络信息资源检索	(50)
第一节 网络信息检索概述	(50)
第二节 网络信息检索技术及发展趋势	(52)
第三节 网络搜索引擎	(55)
第五章 中文工具书检索	(76)
第一节 工具书概述	(76)
第二节 常用检索型工具书	(77)
第三节 常用知识参考型工具书	(97)

第六章 中文图书检索..... (100)

第一节 联机公共目录查询系统 OPAC (100)

第二节 电子图书检索..... (120)

第七章 常用中文期刊信息资源检索..... (144)

第一节 中国知网(CNKI) (144)

第二节 万方数据资源..... (154)

第三节 维普资讯网..... (164)

第四节 人大报刊复印资料系列数据库..... (173)

参考文献..... (179)

第一章 信息的基本知识

在人类发展的历程中,社会形态与经济形态总是同步发展的。经济形态是社会形态的一种表现,每一次人类社会形态的重大发展和进步,都伴随着相应的技术革命的发生,伴随着资源在社会经济中的地位及利用方式的变革,伴随着人类思维方式和生活方式的变革。我们通过对人类社会形态、经济形态发展的研究,不但可以找寻和认识人类发展的轨迹,而且可以更深刻地认识当今的社会形态和经济形态,进而认识和预测人类社会发展的未来。

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。信息社会的发展是一个渐进的过程,有着不同的发展阶段,不可能有一个准确地进入信息社会和信息社会结束的时间。比较公认的观点是:信息社会始于20世纪50年代前后。对其主要标志也有不同的认识,比如:1956年美国历史上从事技术、管理和事务工作的人数第一次超过蓝领工人;1957年前苏联发射了第一颗人造地球卫星;1946年世界上诞生第一台计算机(或1958年IBM公司制成晶体管计算机)。

什么是信息社会?世界各国的理论学家们众说纷纭,一直都没有一个统一的定义,但有一个共同的认识就是:主要依靠信息进行生产的社会为信息社会。因此我们可以将信息社会定义为:在高度工业化之后出现的以计算机、通信、网络等信息技术广泛应用为特征,以信息产业为主导产业,以信息资源的生产、分配和利用为主要目的的社会。在20世纪80年代,关于“信息社会”较为流行的说法是“3C”社会:通讯(communication)、计算机(computer)和控制(control);“3A”社会:工厂自动化(factory automation, FA)、办公自动化(office automation, OA)、家庭自动化(house automation, HA)。到了90年代,关于信息社会的特征又加上多媒体技术、信息高速公路、数字化等。总的来说,信息是信息社会最基本的组成要素,信息社会是一个信息横溢的社会,在飞速发展的信息社会,我们必须充分了解信息的概念和特征。

第一节 信息概述

一、信息的概念

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。那么,究竟什么是信息,什么是知识呢?“信息”一词在中国历史上最早见于唐诗中,拉丁词是 Information,20世纪中叶以后其本质才不断被揭示,并被引入哲学、信息论、系统论、控制论、传播

学、情报学、管理学、通信、计算机科学等领域。信息作为日常用语是指音信、消息。我们每天都身处一个不断地通过感官从外界接受信息的过程。我们阅读,从书刊、报纸中获得信息;我们交流,通过人与人之间的谈话获得信息,等等。但是对于信息的准确概念尚无定论,往往是就信息谈信息。

信息作为一个科学术语,广义的是指事物属性的表征,狭义的是指系统传输和处理的对象,最早是出现于通信领域。在20世纪20年代,哈特莱在探讨信息传输问题时就提出了信息和消息在概念上的差异。实际上任何一种音信和消息或者任何一个系统传输和处理的对象都不外是关于某一事物的某种属性的反映,因此,信息的日常含义与科学含义,广义和狭义都是相通的。

正是由于信息论被广泛地应用于各学科、各领域,因此人们对于信息的认识和定义各有差别。自1948年以来,学术界有关“信息定义”的表述可谓是“千姿百态”,例如:

“信息是用来通信的事实,在观察中得到的数据、新闻和知识。”(《韦伯斯特词典》)

“信息是指对消息接受者来说预先不知道的报道。”(《辞海》)

“信息可以定义为:生物以及具有自动控制系统的机器,通过感觉器官和相应的设备与外界进行交换的一切内容。”(严怡民主编《情报学概念》)

“信息意为消息、情报、知识、资料、数据等。”([美]威尔伯·施拉姆、威廉·波特著《传播学概念》)

“信息就是接受者在接受之前所不知道的知识或消息。”(周海鹏主编《信息技术大辞典》)

总结诸多不同学者的言论,我们认为:信息是被反映事物属性的再现。信息不是事物本身,而是由事物发出的消息、指令、数据等所包含的内容。一切事物,包括自然界和人类社会都会产生信息。

例如:事物本身→发出信息

(1)交通信号灯:红灯停、绿灯行

(2)人的表情(哭或笑):伤心或开心

二、信息的基本特征

1. 客观性

信息是物质间相互作用造成的客观存在,就世界的整体而言,信息统一于物质世界,信息的根源是物质世界,信息的存储、传播依靠物质和能量,它无所谓始,也无所谓终,它与整个物质世界共存。

2. 普遍性

信息是物质的基本属性,信息与物质是辩证的统一。定义中,我们使用了“一切”这个词语,目的是要说明信息是物质的普遍属性和基本属性,信息产生于物质的相互作用,又以物质为载体进行储存和流动;有物质即有相互作用,有相互作用即有信息,信息与物质是共存的,信息并不是只存在于人类社会中。世界是物质的世界,信息是物质固有的普遍属性,信息不仅存在于人类社会,也存在于自然界。人与人之间、机器之间、人与机器之间、动物之间、人与动物之间都是可以进行信息交流的。

3. 测量性 信息代表着物质“客观存在方式……图像集合或符号序列”的一种测量指标或体系,测量通过物质间的相互作用来实现。

4. 流动性(传递性) “物质间相互作用”即导致信息的流动。这里的双重意义是:信息产生于物质间的相互作用,又以物质为载体而实现流动。

5. 中介性 就物质世界的层次来看,信息既区别于物质又区别于精神,它的内核不是具体的物质和能量。尽管有些信息是通过文字、图像等具体的物质形式表现出来的,但它本身却没有质量,也不占有空间。我们见到的占有空间的并不是信息本身,而是存储和携带信息的物质载体。同时,它也不像意识那样依赖于人脑存在,因而不具有主观性,它是介于物质世界和精神世界之间过渡状态的东西,人们是通过它来认识事物的。

6. 增值性 信息是可以累计的。随着事物的不断发展变化,信息也在不断扩充和积累,人们对事物的认识也在不断深入。

7. 可载性 信息虽然是虚体,但信息可以被收集、加工、整理、筛选和存储,通过不同类型的载体实现信息由虚到实的转变。

8. 共享性 信息不具有独一性,它可以多向多次传播,为人们所共享,但不失去其内容,这一点与实物交易不同。

三、信息的形态

形态,即事物的形状和神态。在当代,由于科学技术的发展,信息一般表现为四种形态,即数据、文本、声音、图像。

数据,通常被人们理解为“数字”,这不算错,但不全面。从信息科学的角度来考察,数据是指电子计算机能够生成和处理的所有事实、数字、文字、符号等。当文本、声音、图像在计算机里被简化成“0”和“1”的原始单位时,它们便成了数据。人们储存在数据库里的信息,自然也不仅仅是一些数字。

尽管数据先于电子计算机存在,但是,导致信息经济出现的正是计算机处理数据的这种独特能力。

文本,是指书写的语言——“书面语”,以表示它同“口头语”的区别。从技术上说,口头语言只是声音的一种形式。文本可以用手写,也可以用机器印刷出来。

虽然电子计算机可以代替人们写字,但手写的文字永远具有魅力,不可忽视。在人类目前所处的经济阶段,电子计算机已经学会识别手写的文字,一旦需要,它还能为协议、合同等“验明正身”。

声音是指人们用耳朵听到的信息。在目前的经济阶段,人们听到的基本上是两种信息——说话的声音和音乐。无线电、电话、唱片、录音机等,都是人们用来处理这种信息的工具。

图像是指人们能用眼睛看见的信息。它们可以是黑白的,也可以是彩色的。它们可以是照片,也可以是图画。它们可以是艺术的,也可以是纪实的。它们可以是一些表述或描述、印象或表示——只要能被人们看见就行。

经过扫描的一页文本和数据的图像,也被视为一个单独的图像——虽然新的程序能再次改变这些图像。复印机、传真机、打印机、扫描机是四种不同的,但基本上又是发挥类似功能的机器,所以很可能会在将来的某个时候合而为一。

当然,从技术处理难度来说,在静态的图像和动态的图像、自然的图像和绘制的图像之间,仍存在着很大的差别。

在当代,每一种形态的信息都发生了技术上的重大变化:从大量非立体声到立体声的音乐,从黑白电视到彩色电视,从手拣铅字到电子排版,等等。

同时,文本、数据、声音、图像还能相互转化。一张图画可能相当于 1000 个字,并由 10 万个点组成。“点”又可能是数字、文字或符号。乐谱上的乐曲之所以能被乐师演奏,是因为技术使用者把像点一样的图像转化成了声音;书记录别人口授的语言,则是把声音转变成了文字。当数字化了的信息被输入计算机或从计算机中被输出后,数字又可以用来表示上述这些形态中的任何一种或所有的形态。于是,过去曾被视为毫不相干的行业——计算机、通信、电视、出版等,现在却又成了“亲戚”。

四、信息的功能

功能,即功效和作用。

信息的功能同信息的形态密不可分,并往往融合在一起。打个比方,信息的形态是指信息“是什么模样”,而信息的功能是指信息通过它的形态“能干什么”。

从基本意义上说,信息能通过它的四种形态中的一种形态,“捕捉”到环境中存在的信息——占有它,再把它表示出来,就如同算盘占有了会计师掌握的数字而生成账本一样。同理,打字机占有了作者写出的文字而生成书籍,录音机占有了吉他发出的声音而生成录音带,照片则占有了风景的图像而生成图画。说白了,生成信息就是把已知的信息用一种容易理解的形式发送出去或接收过来。再说白一点,就是把信息数字化,将其整理成“二进制”。

一旦信息被数字化——变成“0”和“1”,所有形态的信息在以后的三种功能中都能加以处理,就好像它们根本就是一码事一样。当照片被分解(“读”)成数字时,图中的每一个点都被赋予一定的值,然后,照片便能通过电话或卫星发送出去或接收过来。数字录音带(DAT)在把声音存进去以后,也要经过类似的处理。

1. 处理信息

处理信息是计算机能为人类作出的一大贡献。计算机能首先进行数据处理,然后进行字处理、声音处理和图像处理。

计算机的处理功能包括转换、编辑、分析、计算和合成。由于利用了半体技术,信息才得以操作和转变信息。虽然今天的计算机已把信息生成、处理和存储功能集于一身,但其处理过程中的各个步骤,就如同在胶片上印上图像那样,彼此是截然不同的:显影、增强、放大,然后把包含在照片上的信息保持在一定的形式中。软件公司通过它们编制的程序,形成了一些处理搜集信息的规则。

2. 储存信息

储存信息通常是指用信息的四种形态中的一种形态来取得信息,并将其保存下来,供日后之用。在古埃及法老时代,文本和数据是储存在古书板上的,而敦煌壁画则储存了我国历代的许多画像。只有声音必须等到工业时代,才能储存在唱片、录音带和激光唱盘之中。而在信息时代,信息则有可能储存在电脑、软盘、光盘之中。

如果储存方式是静态的——只是搜集和保存信息,而没有用信息来做任何事情,这种过程被称为“只读存储”(ROM)。然而,电子时代的储存是动态的。例如,字处理机不但能把人们书写的东西储存起来,而且,在人们需要时还可以进行检索和修改。

信息的生成和储存功能实际上是一件事情的两个方面,而其中关键的就是搜集信息。

3. 传输信息

传输信息是信息的第三种功能。信息传输之所以能够实现,是由于有了电话等手段。在当代有线通信中,传输就是在同轴电缆上用电磁波的速度或在光纤电缆上用光的速度,把各种形态的信息从一端传向另一端。

储存是跨越时间来传输信息,而传输则是跨越空间来传输信息。

简单的传播,诸如利用电话来进行传输,被传输的是声音和图像,而没有将这两者加以改变。然而,当网络不仅传输各种形式的信息,而且也履行生成、处理和储存功能时,便会给正在进行的各种经济活动增加巨大的价值。因此,这样的网络被称为增值网络。

五、信息的类型与载体

一切信息来源于自然界和人类社会。人们的生产、科研、生活以及政治活动等等都是信息产生的源泉,可以说是信息无处不在,但要及时、准确、高效地获取所需信息,就必须充分利用一切可能的信息来源。信息来源按其形式,有口头(语言)型信息源、实物型信息源、传统型信息源、电子型信息源四种类型。

(1) 口头型信息源

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。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,久传易出差错。

(2) 实物型信息源

是以实物如文物、产品样本、模型、雕塑等形式表现的信息资源。秦始皇兵马俑、维纳斯雕像、人造卫星包含着大量信息。其特点是直观性强,感觉实在,信息量大。

口头信息和实物信息目前还没有加工存储和提供检索的工具,使用者只能根据自己的需要去专门地、有目的地搜集、加工、分析、利用。

(3) 传统型信息源

是以文字、图形、符号等各种传统方式记录在传统载体上的知识和信息资源,包括图书、期刊、报纸以及学位论文、专利、标准、会议、政府出版物等特种文献。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。这类信息是当前数量最大、利用率最高的信息资源,是我们获取信息的重要来源。

(4) 电子型信息源

电子型信息源主要指通过使用电子技术实现信息存储和传播的信息源,包括广播、电视、光盘、磁盘、磁带、网络信息源等。网络信息是指因特网上可以利用的各种信息,它包括

科技数据库、新闻、社会科学、文学艺术、学术论文、商业活动、图书、软件等各种信息,具有内容的广泛性、使用的简便性、搜索的网络性和资源的动态性等特点。无论从哪一方面看,网络信息资源几乎胜过了以往所有的传统信息资源,因此,通过网络来获取所需的信息已成为当今获取信息的最主要途径。

信息本身不是实体,必须借助于一定的载体才能表现、传递和利用。载体是信息得以保存的物质实体。从古代的甲骨、金石、棉帛、竹简到现今的纸张、感光材料、磁性材料,信息的载体和存储技术已发生数次质的飞跃,为人类存储、检索和利用信息提供了极大的方便。

当人类跨入 21 世纪这个信息化高度发达的时代,信息同物质、能量一起构成人类社会的三大资源。物质提供材料,能量提供动力,信息提供知识和智慧。因而,信息已成为促进科技、经济和社会发展的新型资源,它不仅有助于我们不断地去揭示客观世界,深化我们对客观世界的科学认识,消除人们在认识上的某种不确定性,而且还源源不断地向人类提供生产知识的原料。

第二节 文献信息资源

一、文献

1. 文献的定义
“文献”一词在中国最早见于孔子的《论语·八佾》一文中,其含义千百年来几经变化:汉代郑玄解释为文章和贤才;宋代朱熹解释为典籍和贤人;宋末元初的马瑞临理解为书本记载的文字资料和口耳相传的言论资料。近现代的一些工具书又将其解释为“具有历史价值的图书文物资料”和“与某一学科有关的重要图书资料”;1983 年我国颁布的《中华人民共和国国家标准:文献著录总则》(GB3792. 1—83)指出,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。”

由上述文献的定义不难看出,文献是由三项基本要素构成的:第一是知识信息内容,这是文献的灵魂所在;第二是载体材料,即可供记录知识或信息的物质材料,如龟甲兽骨、竹木、帛、金石、泥陶、纸张、胶片、胶卷、磁带、光盘等;第三是记录方式,即用文字、图形、代码、符号、声频、视频等方式和技术手段把知识或信息记录在一定物质载体上。知识、载体、记录方式三位一体,不可分割,缺少其中任何一个都不能成为文献。

2. 文献的属性
文献的属性,是指其本身所固有的性质,包括以下四个方面:

(1) 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。

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

(3)人工记录性

文献所包含的知识信息是通过人们用各种方式将其记录在载体上的,而不是天生载于物质实体上的。

(4)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、形式日益多样起来;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。

3. 文献的功能

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。具体功能包含以下三个方面:

(1)存储知识信息

文献是知识的物质存在形式,是积累和保存知识的工具,人类所有的知识成果都只有记录于文献,才能保存和流传;文献的产生是人类文明史上的重要里程碑,人们正是通过文献了解科技信息,通过文献得悉某一科技成果或创造发明诞生于何时、被记录在何种科技文献之中等等具体的情况。

(2)传递知识信息

文献能帮助人们克服时间与空间上的障碍,传递和交流人类已有的知识和经验,促进知识信息的增加和融合,沟通人们思想感情的联系和交流,成为人类知识信息交流的重要途径。

(3)教育和娱乐功能

通过阅读文献,人们可获取科学文化知识,掌握专业技能,提高认识水平和基本素质,还可以娱乐消遣,陶冶情操,丰富精神生活,提高创造能力。

二、资源

1. 资源的概念

所谓的资源是指一切可被人类开发和利用的物质、能量和信息的总称,它广泛地存在于自然界和人类社会中,是一种自然存在物或能够给人类带来财富的财富。对于资源人们又有不同的理解,具体包含的范围也有不同,如:《中国大百科全书》解释为:“广泛存在于自然界的能为人类利用的自然要素。它们是社会生产的原料和燃料的来源,或是社会生产力布局的必要条件和场所。”《英国大百科全书》认为自然资源是:“人类可以利用的自然生成物及生产这些成分的环境的功能,前者如土地、水、大气、岩石、海洋等等,后者如太阳能,地球物理的环境机能(气象、海洋现象、水文地理现象),生态学的环境机能(植物的光合作用、生物的食物链),地球化学的循环机能(地热现象,化石燃料、非金属矿物生产作用等)。”联合国环境规划署对资源的定义是:“所谓资源,特别是自然资源是指在一定时期、地点条件下能够产生经济价值,以提高人类当前和将来福利的自然因素和条件。”

在信息社会,当知识经济条件对某种资源进行利用的时候,必须充分利用科学技术知识来考虑利用资源的层次问题,在对不同种类的资源进行不同层次的利用的时候,又必须考虑地区配置和综合利用的问题。这就是“新资源观”,是在知识经济条件下解决资源问题的认识基础。

2. 资源的类型

从不同的角度对资源有不同的认识,可以将资源区分为不同的类型:

(1)按资源的性质,可以分为自然资源和社会资源

自然资源一般是指一切物质资源和自然过程,通常是指在一定技术经济环境条件下对人类有益的资源,是人类生存和发展的物质基础和社会物质财富的源泉,是可持续发展的重要依据之一。自然资源可以划分为矿产资源、土地资源、生物资源、水资源、气候资源、空气资源等等,也可以分为有形自然资源(如土地、水体等)和无形的自然资源(如光、热资源等)。自然资源具有几个明显的特性:一是有限性:资源的数量与人类社会不断增长的需求相矛盾。二是区域性:资源分布的不平衡,导致存在数量上或质量上的显著地域差异。三是变化性:自然资源在地质条件或人类的影响下不断变化。四是整体性:每个地区的自然资源要素彼此有生态上的联系,形成一个整体。社会资源又可以称之为社会人文资源或社会经济资源,是由人类活动产生的、能直接或间接促进社会和经济发展的物质成果与精神成果的总和,也就是除了自然资源以外的各种资源。社会资源包括人类创造的物质成果,如城市、厂房、农田、机器以及抽象的资金财富等;包括人类本身,即人口资源、人力资源和人才资源;包括人类创造的各种精神财富,如文化、信息、知识、政策、法律、社会秩序等等。其中知识资源既包括掌握知识和技能的人,也包括脱离人脑存在的科学技术(如发明、专利、图纸、文献等)。

自然资源和社会资源都是人类社会和经济活动必不可少的投入,两者的结合才会生产新的社会财富,促进社会与经济的发展。

(2)按资源的再生性,可以分为再生资源和非再生资源

再生资源是指在人类参与下可以重新产生的资源,其更新或再生速度大于或等于我们开发利用的速度。再生资源有两类:一类是可以循环利用的自然资源,如太阳能、空气、雨水、风等等;另一类是生物资源,也称为可更新的自然资源,对于这类资源,应在可能的条件下最大限度地综合利用,充分利用其“可再生”的特点。某些自然资源不能循环再生,或需要漫长的地质时期才能再生,都可称为非再生资源,主要是指矿产资源。再生资源和非再生资源的区分是相对的,如石油、煤炭是非再生资源,但它们却是古生物遗骸在地层中物理、化学的长期作用变化的结果,两者之间可以转化。

(3)按资源利用的可控性程度,可以分为专有资源和共享资源

专有资源是有明确归属和使用权限的资源,如国家控制的资源、管辖内的资源、部门资源等。共享资源是没有归属权、可以共同享用的资源,如阳光、空气、公海、公共信息资源等。

(4)按资源可利用的状况,可以分为现实资源(已经被认识和开发的资源)、潜在资源(尚未被认识,或虽已认识却因技术等条件不具备还不能被开发利用的资源)和废物资源(传统被认为是废物,而由于科学技术的使用,又使其转化为可被开发利用的资源)

(5)按资源的形态,又可分为有形资源和无形资源

有形资源如矿产、海洋、人力资源、设备资源等。无形资源主要是社会资源,如技能、名

誉、环境等。

三、文献信息资源

文献信息资源不仅指印刷型文献及其缩微形式,也包括非印刷型文献(如机读记录、视频记录等)和实物。我们把信息资源中传统文献及其不同的记录形式(电子化、数字化)称为“文献信息”,以区别于其他类型的信息资源。对于文献信息资源的分类,我们着重从以下三个方面来进行划分:

1. 按出版类型来划分

根据出版类型的不同,文献信息资源大体可以分为十三种类型:

(1) 图书(book)

它是通过一定的方法与手段将知识内容以一定的形式和符号,按照一定的体例,记录于一定的载体上,用于保存和传播知识的出版物。联合国教科文组织(UNESCO)的定义是“图书是指不定期的不包括封面在内至少有49页,在某国出版并向公众发行的印刷品”。图书是对已发表的研究成果、生产技术、各种知识和经验系统的总结和整理,有独立的内容体系、相当的篇幅和完整装帧形式,它的特点是内容比较成熟,阐述全面系统,便于我们学习,是系统掌握各学科知识的主要资料。但由于出版周期较长,图书的知识新颖性不够。图书种类很多,包括专著、丛书、教科书、工具书等。

(2) 报纸(newspaper)

它是一种连续出版物,是“以刊登新闻和时事评论为主的定期连续向公众发行的散页出版物”。报纸与期刊不同之处主要是发行周期短,报道速度快,有的一天就有数版;报纸文章的篇幅比较短小,内容多以新闻、动态报道为主,研究性文章很少,有些知识性报纸主要刊登面向大众的普及知识。报纸对于决策者、经营管理者、研究人员等都是非常重要的信息源。

(3) 期刊(journal)

又称杂志,也是一种连续出版物,一般是指具有固定题名,统一出版形式,定期或不定期出版的连续出版物。其特点是出版周期短,报道文献速度快,内容新颖,发行及影响面广,能及时反映科学技术和社会生活中的新成果、新水平、新动向。期刊发表的论文大多数是原始文献,许多新成果、新观点、新方法往往首先在期刊上刊登,因此有很高的参考价值。但期刊论文由于时效性较强,需要及时查找和利用。期刊论文是文献信息资源的主要类型之一,是检索工具报道的主要对象。它按照内容性质可分为学术性期刊、通报性期刊、技术性期刊、科普性期刊、动态性期刊、综述与述评性期刊和检索性期刊等类型。

(4) 报告(report)

报告是政府、科研、企业等部门围绕某一专题进行研究取得成果的总结报告,或是研究过程中的阶段进展报告。报告的类型有技术报告(technical report)、札记(notes)、论文(paper)、备忘录(memorandum)、通报(bulletin)、可行性报告(feasibility report)、市场预测报告(market prediction report)、进展报告(progress report)、初步报告(primary report)、中间报告(interim report)、最终报告(final report)等等。报告的出版特点是各篇单独成册,统一编号(报告号),由主管机构连续出版。报告一般单独成册,有具体的篇名、机构名称和统一的连续编号。

报告根据内容一般可以划分为绝密(top secret)、秘密(secret)、保密(classified)、非密

限制发行(restricted)、解密(declassified)及公开(unclassified)等几种密级。报告在内容上比期刊论文等专深、详尽、可靠,能反映一个国家、某一学科领域、某一行业或某一部门的科研水平,是一种重要而又不易获取的信息源。如PB(Publishing Board)报告、AD(Astia Documents)报告、NASA(National Aeronautics and Space Administration)报告和DOE(Department of Energy)报告是著名的美国政府的四大报告。

(5) 专利文献(patents)

专利是发明人创造发明了某种新技术,通过一定的申请程序,经政府专利局审批后,即获得一定年限(时期)的垄断权。专利分为发明专利、实用新型专利、外观设计专利,是一种知识产权,在专利有效期内可以交换、继承、转让。我们通常所说的专利有三种含义:一是指专利权;二是指受到专利法保护的发明创造;三是指专利文献。

专利文献是各国专利局以及国际性专利组织,在审批专利过程中产生的官方文件及其出版物的总称,包括发明专利申请书和发明专利说明书,也包括有关发明的其他类别的文件,还包括专利局公开出版的各种检索工具书(如专利年度索引、专利公报等)。我国出版的专利文献主要有:发明专利公报、实用新型专利公报和外观设计专利公报;发明专利申请公开说明书、发明专利说明书;实用新型专利说明书;专利年度索引。从检索和利用的角度来说,专利文献主要是指专利说明书,是专利申请人向专利局递送的说明其发明创造的文件。专利说明书的内容比较具体,有的还有附图,通过它可以了解该项专利的主要技术内容。专利公报是各国专利局或国际性专利组织报道专利申请审批状况及相关法律、法规信息的定期出版物,报道一定期间内专利申请的公开、公告、专利授权、申请撤销、驳回或专利失效的有关信息,多为周刊。因此,可以说专利文献已成了获取科技信息重要的来源。

(6) 会议文献(conference paper)

会议文献是指国际学术会议和各国国内重要学术会议上发表的论文和报告。按会议的级别和规模可分为国际会议文献、全国会议文献和地区性会议文献。会议文献通常分为会前文献和会后文献两类。前者主要是指论文预印本和论文摘要,后者是指会议的论文汇编。会议文献除了以会议录的形式出版外,也有不少会议文献在期刊上发表。会议文献学术性强、内容新颖、质量较高,往往能代表某一领域的最新的研究成果及水平,从中可了解国内外科技发展趋势,有较大的参考价值,是重要的信息来源。

(7) 标准文献(standards)

标准文献是指由技术标准、管理标准、工作标准及其他具有标准性质的类似文件所组成的一种特种文献体系,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。它是在有关方面的合作下,按照规定程序编制并经主管机构批准,以特定形式发布,供一定范围内广泛而多次使用的,包括一整套在特定活动领域内必须执行的规格、定额、规则、要求的文件。它要与现代科学技术和生产发展水平相适应,并且随着标准化对象的变化而不断补充、修订、更新换代。经权威机构批准的规章性文献,具有一定的法律约束力。

从检索和利用角度来说,标准主要是指技术标准(technical standards),一种规范性的技术文件,它是在生产或科学研究活动中对产品、工程或其他技术项目的质量品种、检验方法及技术要求所作的统一规定,供人们遵守和使用。技术标准按其使用的范围可分为:国际标准、区域性标准、国家标准、专业标准和企业标准等五个级别。

标准文献的特征是有统一的产生过程,有明确的适用范围和用途,编排格式、叙述方法

严格统一,并且有统一的代号和编号。同时,还具有可靠性、现实性、协调性、系统性、完整性、时效性和法规性等特点。一个国家的标准文献反映该国的生产工艺水平和技术经济政策,而国际现行标准则代表了当前世界水平。国际标准和工业先进国家的标准是科研生产活动的重要依据和信息来源。

(8) 学位论文(thesis dissertation)

学位论文是高等学校、科研机构的研究者为获得学位,在进行科学研究后撰写的学术论文。学位论文的质量参差不齐,但都是就某一专题进行全面的文献调查,比较详细地总结前人的工作和当前的研究水平,并作系统的实验研究及理论分析,一般都是具有独创性,其数据较全、探索较深,并且还附有大量的参考文献,是一种重要的文献来源。学位论文一般来说是不公开出版的,仅由学位授予单位和国家指定单位收藏。我国法定的学位论文收藏单位有三个:国家图书馆(收藏所有文理科博士学位论文及博士后科技报告);中国科技信息研究所(收藏理工科的硕士学位论文);中国社科院文献中心(收藏文科及语言科的硕士论文)。学位论文中也有极少数会以科技报告、期刊论文的形式发表。

(9) 技术档案(technical records)

技术档案是指生产建设、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、设计图表、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。

(10) 政府出版物(government documents)

政府出版物是指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。它所包括的内容范围广泛,几乎涉及整个知识领域,既有基础科学、应用科学方面的,也有政治、社会、经济、文化等社会科学方面的。政府出版物可分为行政性文件(如政策、法规)和科技文献(科技报告、技术档案)两大类。某些国家的政府出版物习惯使用某种颜色的封面来称呼,如白皮书、蓝皮书等等。

政府出版物内容可靠,有极高的权威性。通过政府出版物,可以了解某一国家的科技政策、经济政策、法令、规章制度、科技活动、科技成果等。

(11) 产品资料(product literature)

产品资料主要是指产品目录、产品样本和产品说明书等厂商用来对产品进行宣传 and 使用的资料。产品样本通常对定型产品的性能、构造、原理、用途、使用方法和操作方法、产品规格等作出具体说明。产品样本在技术上比较成熟,数据比较可靠。搜集和分析产品样本,有助于了解某类产品的水平、现状和发展动向,对于新产品的设计、试制,对于产品的选购、使用等有着较大的参考价值。

(12) 电子预印本文献(e-print)

也称为电子版的学术文献,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 preprint 和 postprint 两种形态的文献。前者是指论文原稿完成后,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;后者是指经过同行评审,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,preprint 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、费用低等特点。在 1991