

高 等 学 校 教 材

王园春 李瑞斌◎主 编 | 韩 颖◎主 审

科技信息 检索与利用

KEJI XINXI
JIAN SUO YU LIYONG

石油工业出版社



教育部 高等学校计算机专业类课程 教材系列

科技信息 检索与利用

Paul R. Wilson

美国威斯康星大学威斯康星分校教授

清华大学出版社

G252.7/93

2006

高等学校教材

科技信息检索与利用

王园春 李瑞斌 主编

韩 颖 主审

石油工业出版社

内 容 提 要

本书是根据高等院校理工科学生信息教育课程的要求而编写的教材。全书共分为十章,主要介绍了信息检索基础知识、国内外信息的手工检索及网上查询、特种文献及网上检索、信息检索与咨询服务及文献写作等内容。

本书可作为高等院校“信息检索与利用”课程的教材,也可作为广大信息用户及科研人员进行信息检索的指导书。

图书在版编目(CIP)数据

科技信息检索与利用/王园春,李瑞斌主编.

北京:石油工业出版社,2006.12

(高等学校教材)

ISBN 7-5021-5621-6

I. 科…

II. ①王…②李…

III. 科技情报-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(2006)第 116847 号

出版发行:石油工业出版社

(北京安定门外安华里 2 区 1 号 100011)

网 址:www.petropub.cn

发行部:(010)64210392

经 销:全国新华书店

排 版:北京乘设伟业科技排版中心

印 刷:石油工业出版社印刷厂

2006 年 12 月第 1 版 2006 年 12 月第 1 次印刷

787 × 1092 毫米 开本:1/16 印张:12.25

字数:310 千字 印数:1—6000 册

定价:22.00 元

(如出现印装质量问题,我社发行部负责调换)

版权所有,翻印必究

前 言

21 世纪社会信息化已成为不可阻挡的历史潮流,在这个时代,科学、技术、经济的竞争都将以信息的竞争为前奏,信息用户教育必将成为高等教育中越来越重要的内容。

20 世纪 80 年代初,各高等院校相继开设了具有中国特色的“科技文献检索与利用”课程,相关教材也相继出现。随着计算机、网络通信技术的高速发展及国际因特网的建成,信息的环境发生了变化,人们获取信息的方式和手段也发生了根本变化。因此培养高等院校的学生应用现代化技术手段,特别是网络技术来获取各种信息、知识已成为当前信息用户教育的主要内容。为适应现代信息技术的不断发展,满足现阶段高等院校信息用户教育的教学需要,结合多年的教学实践经验,我们编写了此教材。

本教材具有以下几个特色:

第一,注重理论基础知识的阐述。本教材用较大篇幅对文献信息及检索的诸多环节及其相互关系进行了详细的论述,用规律性的原理对检索过程进行理论上的指导,以培养学生“举一反三”的能力,从知识创新、信息意识、信息素质教育等角度强调了信息检索课程的重要性。

第二,对原教材网络检索与应用的有关章节进行了修编,充实了网络信息检索方法与技巧的内容以及各种应用数据库的使用方法,并对国际、国内著名的检索系统的光盘版和网络版都作了图文并茂的介绍,从而把传统的文献检索扩展到现代化的信息查询。

第三,从工程技术、化学化工、电气及电子工程、石油等专业方面详细总结和介绍了相关信息的手工检索工具及网络检索方法。

本书第一章、第四章、第五章、第六章、第八章由王园春编写;第二章、第七章由李瑞斌编写;第三章由孙宁、张美薇编写;第九章由王园春、李瑞斌编写;第十章由孙宁、张美薇编写。

本书由韩颖任主审,由王园春、李瑞斌任主编,由孙宁、张美薇任副主编。本书在编写过程中得到了王文广教授和金秋颖主任的指导,在此致以诚挚的谢意。

限于编者的水平有限,书中难免有瑕疵之处,敬请各位专家和读者批评指正,以便修改完善。

编 者
2006 年 10 月

目 录

第一章 信息检索基础知识	(1)
第一节 信息、知识与情报	(1)
第二节 信息源及其种类	(5)
第三节 知识创新、信息意识与信息素质教育	(13)
第四节 信息检索概述	(18)
第五节 信息检索工具	(22)
第六节 信息检索语言	(24)
第七节 信息检索程序、策略及效果评价	(31)
第二章 现代信息检索基础知识	(35)
第一节 计算机信息检索技术及策略	(35)
第二节 数据库及其种类	(40)
第三章 中文科技信息的综合性检索	(44)
第一节 国内检索工具概况	(44)
第二节 中文科技信息的手工检索工具	(45)
第三节 CNKI 系列数据库	(47)
第四节 中文科技期刊数据库	(55)
第五节 万方数据资源系统	(58)
第四章 外文科技信息的综合性检索	(63)
第一节 美国《科学引文索引》	(63)
第二节 前苏联《文摘杂志》	(65)
第三节 日本《科学技术文献速报》	(67)
第四节 外文综合性科技信息数据库	(70)
第五章 国内外专业信息检索与利用	(78)
第一节 工程技术类信息的检索	(78)
第二节 化学、化工信息检索	(83)
第三节 电气、电子工程信息检索	(91)
第四节 石油信息检索	(94)
第六章 特种文献检索	(104)
第一节 专利文献及其检索	(104)
第二节 标准文献及其检索	(112)
第三节 会议文献及其检索	(116)
第四节 科技报告及其检索	(119)

第七章 计算机检索系统及利用	(122)
第一节 重要国际联机检索系统	(122)
第二节 光盘检索系统	(129)
第八章 网络信息检索与利用	(139)
第一节 网络信息检索概述	(139)
第二节 因特网与万维网	(142)
第三节 搜索引擎	(144)
第四节 因特网检索方法	(148)
第九章 文献信息的利用	(152)
第一节 文献信息的收集、整理和分析	(152)
第二节 科技文献写作的意义及步骤	(157)
第三节 科技综述与科技述评的写作方法	(160)
第四节 学术论文的写作方法	(162)
第五节 开题报告与结题报告的写作方法	(170)
第六节 毕业论文的写作方法	(174)
第十章 信息检索与咨询服务	(181)
第一节 解答咨询服务	(181)
第二节 定题服务	(182)
第三节 科技查新服务	(184)
第四节 网络信息资源查询辅导服务	(185)
参考文献	(188)

第一章 信息检索基础知识

伴随着人类进入 21 世纪,世界充满着取之不尽、用之不竭的信息。现代信息技术迅猛发展,信息已成为人类社会发展的—种驱动力,人们越来越重视对信息资源的有效开发与利用。信息是一种极其重要的社会财富,信息同物质、能量构成人类社会的三大重要战略资源。物质提供材料,能量提供动力,信息提供知识与智慧。因此,信息已成为促进科技、社会、经济发展的新型资源,它不仅有助于人们不断地揭示客观世界,深化人们对客观世界的科学认识,消除人们在认识上的某种不定性,而且还源源不断地向人类提供生产知识的原料。

随着科学技术的进步,尤其是自因特网问世以来,信息无节制地膨胀。信息的大爆炸与人们有限的时间、精力和吸收能力产生了巨大的剪刀差。如何快捷、准确、经济地检索所需要的信息,缩小甚至消除这种剪刀差,就成为当前急需解决的问题。

第一节 信息、知识与情报

一、信息

1. 信息的概念

信息这一概念首先是由 Shannon 和 Weaver 提出来的,他们试图给信息一个正式和定量的定义,从通信工程、计算机和电信的角度出发,把一个消息中携带的信息量的大小用节(bit)来度量。

信息作为一个科学术语最早出现于通信领域,20 世纪中叶后被引入哲学、信息论、系统论、控制论、情报学、经济学、管理学、计算机等领域。时至今日,人类正进入信息时代,信息无处不存在,但对信息的定义却众说纷纭,不同的学者从不同的研究角度和使用目的出发,有着不同的见解。信息的英文词汇是 Information,目前对这一词汇的解释也有 100 余种。

美国学者威尔伯·施拉姆在其所著的《传播学概念》中给信息下的定义为:信息意为消息、情报、知识、资料、数据等。

我国情报专家严怡民在其主编的《情报学概念》一书中说,信息可定义为: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。

哲学家从产生信息的客体来定义信息,认为信息是事物本质、特征、运动规律的反映。

经济学家认为,信息是与物质、能量相并列的客观世界的三大要素之一,是为管理和决策提供依据的有效数据。

心理学专家认为,信息是存在于意识之外的东西,它存在于自然界、印刷品、硬盘以及空气之中。

信息论的创始人申农(C. E. Shannon)从通信系统理论的角度把信息定义为:用来减少随机不确定性的东西。他主要研究信息的本质,用数学的方法研究信息的计量、传递、变化、储存,其核心是信息熵的概念。信息熵就是衡量对某种事物认识的不确定度,即信息的有序度。

申农从这一思想出发推导出了信息测度数学公式,为信息认识做出了重大的贡献,也标志着信息科学进入了定量研究阶段。

控制论专家 N·维纳认为,信息就是我们适应外部世界,并把这种适应反作用于外部世界的过程中,同外部世界进行相互联系、相互交换的一种内容;《辞海》中对信息的定义为:信息是指对消息接受者来说预先不知道的报导。

在新闻界,信息被认为是对事物运动状态的陈述,是物与物、物与人、人与人之间的特征传输。

图书信息领域的专家又把信息视为可以以各种形式进行传播、记录、出版及发行的观念、事实及论著。

美国《韦氏字典》把信息定义为:信息是用来通信的事实,在观测中得到的数据、新闻和知识。

以上各学科的学者、专家以及有关领域对信息的这些定义都是从信息的受体、内涵和控制论等角度对信息的属性所作的描述。因此,我们说信息是事物属性的再现。信息不是事物本身,而是由事物发出的、体现它存在和运动状态的信号、消息、指令和数据等所包含的内容。即广义的信息可定义为:信息是事物属性的表征;而狭义的信息则可定义为:信息是指系统传输和处理的对象。

资讯与信息的涵义是相同的,特别是那些通过电子方式记录、传播的信息。资讯一词是海外学者对 Information 的译名。

2. 信息的属性

信息的属性主要有:

(1)信息具有客观性与普遍性。世界上一切事物都是在运动中,都有一定的运动状态和状态方式的改变,因而一切事物随时都在产生信息,即信息的产生源于事物,是客观事物普遍性的表征,它可以被感知和共享、被处理和存储、被传递和利用。

(2)信息具有多样性与综合性。信息在不同的领域具有多种不同的特性或表现形式,如客观事物中的各种自然属性、人工设备的技术特征、人类社会的各种社会特征、人脑中反映客观事物认识的思想、知识、人类交流信息过程中的声音、文字、图像以及用各种编码形式记录下来的数据、新闻、情报和消息等。各种形式的信息又常常以综合的方式表现事物的特征,所谓“多媒体”正是信息多样性和综合性的集中表现。

(3)信息具有流动性与传递性。信息在事物之间的相互联系必定在信息的流动中发生。信息的传递性表现在人与人之间的消息交换,人与自动机、自动机与自动机之间的信息交换,动物界和植物界的信号交换。同时,人类进化过程中的细胞、遗传也被看做是信息的传递与交换。

(4)信息具有相对性与有效性。从信息作为事物相互联系的反映角度看,信息源不确定的程度或者信息源接受信息量的多少,均与信宿的状态有关。这一特征在人作为信宿接受信息的过程中表现得尤为明显。同一信息对具有不同认知水平的人所产生的作用和有效性也不同。

(5)信息具有积累性与价值性。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,但信息只有被利用才会产生价值,否则其价值会随时间的流逝而减少,或成为“信息垃圾”。

(6)信息具有中介性与共享性。信息源于事物,但不是事物本身,是人们用来认识事物的媒介。信息能够共享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以上的用户分享,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之而损失或减少。

3. 信息的类型

在人类社会和自然界里,不同领域对信息有不同的定义范围,根据不同的定义范围,从不同的角度来划分信息的种类。

(1)按信息产生的客体性质来划分,可分为自然信息与社会信息。

(2)按信息的社会属性划分,可分为政治信息、军事信息、经济信息、科技信息、管理信息和生活信息等。

(3)按信息传递范围划分,可分为公开信息、半公开信息和非公开信息。

(4)按信息的加工程度划分,可分为零次信息、一次信息、二次信息、三次信息和高次信息。

(5)按价值观念划分,可分为有价值信息和无价值信息。

(6)按运动状态划分,可分为动态信息和静态信息。

(7)按信息所依附的载体形式划分,可分为文献信息、口头信息、电子信息等。

二、知识

1. 知识的概念

在《辞海》中把知识定义为:人们在实践中积累起来的经验,从本质上说,知识属认识范畴。在《现代汉语词典》中对知识的解释是:知识是人们在改造世界的实践中所获得的认识和经验的总和。这两种解释均把知识作为实践或实践活动的认知成果来看待,而知识的本质则是认知活动中的主体与客体的动态关系。

知识是人类在认识和改造世界的社会实践中获得的对事物本质认识的成果和结晶,是人的主观世界对于客观世界的概况和如实反映。知识是人类通过信息对自然界、人类社会以及思维方式与运动规律的认识,并通过人的大脑进行思维重新整合,使信息系统化从而构成知识。因此,人类不仅要通过信息感知世界、认识世界和改造世界,而且还要根据所获得的信息组成知识。由此可见,知识是信息的一部分。

知识虽与信息密不可分,但知识并不等于信息。人们为了进行知识的传递和交流,必须使知识具有能为感觉器官所感知的形式,即借助于文字、语言、符号、代码、电磁波、图像和实物等加以表现。这种表现形式就是信息。因此,只有将反映自然现象和社会现象的信息经过加工,上升为对自然和社会发展客观规律的认识,这种再生信息才构成知识。正像英国著名情报学家 B·C·布鲁克斯对信息与知识关系的描述,即信息是使人原有的知识结构发生变化的那一小部分知识。

知识在现代社会中起着巨大的作用,知识是文明程度的标志。衡量一个国家、一个民族文明程度的高低,主要看其创造、吸收、掌握、应用知识的能力。在知识经济时代,知识已经成为一个庞大的、社会化的产业,它渗透于人类活动的各个领域。同时,知识又是建设精神文明的动力,是科学教育的内容,能促进人类智能的改善。

2. 知识的属性

(1) 知识具有意识性。知识是一种观念形态的东西,只有人的大脑才能产生它、认识它、利用它,知识通常以概念、判断、推理、假说和预见等思维方式和范畴体系表现自身的存在。

(2) 知识具有信息性。信息是产生知识的原料,知识是经人类认识、理解并经思维重新整合后的系统化信息,知识是信息中的一部分。

(3) 知识具有实践性。社会实践是一切知识产生的基础,也是检验知识的标准,科学知识对实践有重大的指导作用。

(4) 知识具有规律性。人们对实践的认识,是一个无限的过程,人们在这种无限的过程中所获得的知识从一定的层面上揭示了事物及其运动过程的规律性。

(5) 知识具有继承性。每一次新知识的产生,既是原有知识的继承与利用、深化与发展,又是更新知识产生的基础与前提。知识被记录或物化为劳动产品后,可以世代相传并利用。

(6) 知识具有渗透性。随着人类认识世界的不断深化,各种门类的知识可以互相渗透,形成了许多新的知识门类,构成了科学知识的网状结构体系。

3. 知识的类型

(1) 按知识反映的对象划分,可分为生活常识和科学知识。

(2) 按知识成熟的程度划分,可分为经验知识和理论知识。

(3) 按知识存在的方式划分,可分为主观知识和客观知识。

(4) 按知识的门类、结构划分,可分为基础知识、技术知识和应用知识。

(5) 按知识描述的内容划分,可分为哲学知识、自然科学知识、社会科学知识和思维科学知识等。

三、情报

1. 情报的概念

前苏联情报学家米哈依诺夫认为:情报是作为存储、传递和转换对象的知识。我国著名科学家钱学森曾经说过:情报就是为了解决一个特定的问题所需要的知识。我国情报界近年来提出:情报就是一种信息;情报,即为一定目的、具有一定时效和对象,传递着的信息等。因此,情报就是人们在一定的时间内为一定的目的而传递着的有使用价值的知识或信息。情报是一种普遍存在的社会现象,人们在社会实践中源源不断地创造、交流与利用各种各样的情报。

2. 情报的属性

(1) 情报具有知识性。情报的知识性是指情报的本质是知识。在日常工作和生活中,我们通过各种媒体,随时都在接受、传递和利用大量的感性和理性的知识。这些知识中就包含着我们所需的情报。反过来说,没有一定的知识内容,任何东西都不能成为情报。

(2) 情报具有传递性。情报的传递性是指无论多么重要的知识,人们不知道其存在就不能成为情报。知识要变成情报,还必须经过运动。我国著名科学家钱学森曾经说过,情报是激活的知识,即人们通过主动搜集情报,促使静态知识成为动态情报。

(3) 情报具有效用性。情报的效用性是指运动着的知识也不都是情报,只有那些能满足特定需要的运动的知识才可称之为情报。

3. 情报的类型

(1) 按情报的内容范围划分,可分为技术情报、社会科学情报、政治情报、军事情报、经济情报、技术经济情报、体育情报和管理情报等。

(2)按情报的使用目的划分,可分为战略情报和战术情报。

(3)按情报传播的形式划分,可分为口头情报、实物情报、文献情报以及文字情报、数据情报和音像情报等。

(4)按情报公开的程度划分,可分为公开情报、内部情报、秘密情报和机要情报等。

第二节 信息源及其种类

一、信息源

信息源,顾名思义,就是信息的来源。在国内图书情报领域,信息源被解释为人们在科研活动、生产经营活动和其他一切活动中所产生的成果和各种原始记录,以及对这些成果和原始记录加工整理得到的成品。信息源可划分为文献信息源和非文献信息源。根据对信息源的加工层次和集约程度,信息源可划分为四个层次:所有物质均为一次信息源,也称本体论信息源,从一次信息源中提取信息是信息资源生产者的任务;二次信息源也称感知信息源,主要储存于人的大脑中,传播、咨询、决策等领域所依赖的主要是二次信息源;三次信息源又称再生信息源,主要包括口头信息源、体语信息源、文献信息源、电子信息源和实物信息源,其中又以文献信息源(包括印刷型和电子型文献信息源)最为常用;高次信息源也称集约信息源,是文献信息源和实物信息源的集约化和系统化,前者如档案馆、图书馆、数据库,后者如博物馆、样品室、展览馆、标本室等。

对信息源的划分角度不同,就产生很多种类。

1. 按信息的表现形式划分

按信息的表现形式可分为以下几种:

(1)文献型信息源。该种信息源是存储语言文字形式信息的各种载体的集合。文献型信息源是目前信息内容最丰富、可靠的信息,是人们使用最多的信息源。

(2)非文献型信息源。其中包括电子信息源(数值型信息源、声像型信息源、多媒体信息源)和实物及口头信息源。

① 数值型信息源,即存储数据形式信息的载体的集合。

② 声像型信息源,即存储声音或图像信息的载体的集合,如磁带、广播、电视。

③ 多媒体信息源,是一种时代发展的产物,它集文字、声音、图像于一体,多以光盘或因特网(Internet)上资源的形式出现,是目前发展最快、数量最多的一种信息源。

④ 实物及口头信息源。实物信息源是指自然实物和人工实物中所含信息的集合,口头信息源是指在交流、讨论、报告过程中所含的信息集合。

2. 按信息内容的加工程度划分

按信息内容的加工程度可分为零次信息源、一次信息源、二次信息源、三次信息源和高次信息源。

3. 按信息内容的保密程度划分

按信息内容的保密程度可分为公开信息源、半公开信息源和非公开信息源。

4. 按信息内容的应用范围划分

按信息内容的应用范围可分为科技信息源、经济信息源、管理信息源及生活信息源。

5. 按信息的载体和传输形式划分

(1)按载体形式划分可分为:

- ① 印刷型,即以纸张为介质,以铅印、油印、胶印、复印等手段记录信息的载体。
- ② 缩微型,即以感光材料为介质,以缩微照相为手段记录信息的载体。
- ③ 机读型,即以磁性材料为载体,通过编码和程序设计,由计算机输入和输出的信息。
- ④ 声像型,即以电磁材料为载体,借助特殊设备,直接将声音和图像等信息记录下来的一种动态信息。

(2)按信息传输形式划分可分为:

- ① 网络信息源,即各种网络上的信息集合。
- ② 非网络信息源,即不用通信设施就能获得的信息集合。

二、文献信息源

1. 文献的含义

随着社会生产力的发展,出现了各种不同的表达人们思想的物质载体,发明了各种各样记录信息的方式。因此,文字也不再是表达思想的唯一手段,而图形、符号、声频、视频同样成为人们表达思想、传递信息的重要手段。

在《文献情报术语国际标准(草案)》(ISO/DIS 5127)中对文献是这样定义的,即为了把人类知识传播开来和继承下去,人们用文字、图形、符号、声频、视频等手段将其记录下来,或写在纸上,或晒在蓝图上,或摄制在感光片上,或录到唱片上,或存储在磁盘上。这种附着在各种载体上的记录统称为文献。在 GB 3792. 1—1983《文献著录总则》中,对文献是这样定义的,即文献,记录有知识的一切载体。

权威部门对文献颁布的这两种定义揭示了文献的特征。由此可知,文献由四要素构成,即文献信息、文献载体、符号系统和记录方式。四位一体不可分割,缺少任何一个都不能构成文献。

2. 文献的特点

现代科技文献具有以下特点:

(1)文献数量急剧增长。随着科技的发展、科研人员数量的增加和日益庞大的信息系统的出现,使得科技成果以前所未有的速度增长。据不完全统计,世界每年出版图书 80 万种以上,发表科技论文 600 万余篇,科研报告约 70 万余件,发表专利说明书约 100 万余件。不同学科,文献量增长速度也不相同。科技文献比哲学、社会科学文献增长的速度快,前者尤以化学、化工文献增长为最快。《美国化学文摘(CA)》历年收录的信息条数可以说明化学、化工文献增长的速度(见表 1-1)。

表 1-1 《CA》历年收录的信息条数

年份	收录条数,条	年份	收录条数,条
1907	7975	1977	409841
1917	15601	1987	477177
1927	32909	1997	716564
1937	36038	1998	671008
1947	38386	1999	720545
1957	101027	2000	725195
1967	243982		

此外还有大量的特种文献出版,尤其是随着因特网的快速发展,网上传播的信息量非常庞大。

(2)文献内容交叉重复。文献内容交叉重复表现在科技文献经常由一种类型转化为另一种类型重复发表。同时,语种繁多、译文增加,也造成了科技文献在内容上的交叉重复现象,造成人力物力的浪费。但是在一定条件下,也为获取原始文献带来一定方便。

(3)文献出版分散。现代科技文献出版分散现象表现在两个方面:一是同一专业的文献往往分散刊载在许多相关专业或综合性刊物上;二是专业性刊物所发表的文献一般涉及多种科学领域的科研成果,表现了相关或相近学科相互交叉渗透的特点。文献这种分散出版现象也给文献的利用带来不便。

(4)文献失效加快。社会的进步、科技的发展,时时刻刻都有新的发明创造,使得科技文献也随之出现新陈代谢加快、老化加剧、有效使用寿命日益缩短的趋势。文献学家贝尔(J. Bernal)、保尔登(R. Barton)和凯布勒(R. Kabler)先后提出了文献老化的半生期(Half-life),用半生期的概念来解释某学科信息的老化速度及使用寿命,即某学科现时尚在利用的全部文献中的一半,是在多长时间内发表的。例如,化学文献的半生期为8.1年,即现正在利用的化学文献的50%,其出版年限不超过8.1年。文献的半生期越短,说明其信息的老化速度越快,使用寿命越短。文献的老化速度与科学信息量的增长有关。有关学科文献的半生期见表1-2。

表1-2 有关学科文献的半生期

学科名称	半生期,年	学科名称	半生期,年
生物医学	3	化学	8.1
冶金学	3.9	植物学	10
物理学	4.6	数学	10.5
化工	4.8	地质学	11.8
机械制造	5.2	地理学	16

(5)文献载体类型多样化。计算机技术和现代信息存储技术的应用,使文献的载体从纸媒介向光学、磁性媒介发展,文献的缩微化、电子化已成为主要的发展趋势。因此,出现了印刷品与缩微资料、声像资料、机读资料、光盘资料等多种类型文献并存的趋势。其中光盘以海量存储著称,配以多媒体技术,发展尤为迅速。电子网络版以数据库和电信网络为基础,以计算机的硬盘为载体。人们还可以利用因特网这一分布全球的因特网络,多途径、多选择、多层次地检索所需信息。

(6)文献语种增多。科技文献语种的多样化,已成为读者利用科技文献的一大障碍。据报道,科技文献出版的文种有80余种之多,比较集中的文种分布也不下10种。文种的多样化严重地影响了科技文献的收集、整理、检索和利用。

3. 文献的构成

文献由四要素构成: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,符号系统是信息的携带者,载体是符号赖以依附的“寄主”,而记录方式则是代表文献的符号进入载体的方法和过程,四要素缺少任何一种都不可能形成文献。

(1)文献信息。文献信息有以下特点:文献信息是经人类加工后记录下来的信息;文献信息的传递是通过人工符号系统——文字、标识符、声像信号来实现的,因而对文献信息的摄取

方式和吸收的程度必然受到这种人工符号的制约;文献信息是一种相对固化的信息,如印刷品所传递的信息是无法变动的,可擦写光盘和磁盘虽然可以用修改、删除、增入等手段更新数据,一旦变动后又处于静态之中;文献所传递的信息是人对客观世界的反映,因而不一定完全符合客观世界表现出的信息内容,这种“歪曲”、“变形”、“失真”的程度因人而异。这些特点告诉我们,信息资源的开发涉及载体资源的开发。

(2) 文献载体。与其他信息载体不同,文献载体必须适应于文献符号和相应的记录方式,同时又要有利于传播、整理和长期保存。文献载体的材质随着人类文明的演进不断地进行新陈代谢,从天然的泥版、岩石、兽骨、木版、竹片、兽皮、树叶、桦树皮等到各种人工材质,如无机材质(陶、砖瓦、瓷、玻璃等)、高分子材质(纸、胶片、醋酸纤维等)、复合材质(磁带、激光盘等)。

(3) 符号系统。文献中的符号系统是指图画、文字、公式、图表、编码、声像和电磁信息等。

(4) 记录方式。文献记录方式是指将表达信息的符号系统通过特定的人工记录方式使其附着于一定的文献载体上。记录方式的种类很多,按记录手段分,有写画、雕刻、印刷、摄制、录音;按记录方法分,有手工记录、机械记录、光记录、电记录、声记录和磁记录。

4. 文献的种类

(1) 按出版类型划分,文献信息主要有以下几种:

① 科技图书。科技图书大多是对已发表的科研成果、生产技术和经验或者基本知识领域系统的论述或概括,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、组织而编写成的。不少科技图书的内容还包含一些从未发表过的研究成果或资料。

科技图书的特点:内容比较系统、全面、成熟、可靠,有一定的新颖性;但编辑出版时间过长,传递信息的速度太慢,包含的内容一般只是反映3~5年以前的研究水平。

科技图书是综合、积累和传递科技知识,教育和培养科技人才的一种重要工具,它可以帮助人们比较全面、系统地了解一特定领域中的历史和现状,可以将人们正确地带入自己所不熟悉的领域,还可以作为一种经常性的查考工具。从信息检索角度来看,科技图书一般不作为主要检索对象,研究人员利用图书的比重比较小。美国有的信息专家曾经对美国各大学的科学家和英国电气工程师们进行过调查,发现在他们所阅读的各种科技文献中,图书的比重分别占19%和14%。

② 科技期刊。科技期刊是一种以印刷形式或其他形式逐次刊行的,通常有数字或年、月顺序编号的,并打算无限期地连续出版下去的出版物(ISO 3297—1986)。

广义的期刊则包括一切定期刊行或不定期刊行的连续性出版物,如杂志、报纸、年度报告、年鉴、丛书以及学会的会议录、学报和纪要等。

科技期刊在科学技术活动中一直起着非常重要的作用,是科学交流的主要工具。科技期刊具有以下特点:数量大、品种多、内容丰富多样;出版周期短、报道速度较快;发行、流通面很广泛;连续性强;伴随着相应的学科领域发展而发展。

③ 会议文献。会议文献是指在国内外各种学术会议上交流的论文,以及由此汇编成册或内部交流或公开出版的文献。会议文献按出版时间可分为会前出版物(包括会议日程表、会议议程、会议论文预印本和论文摘要等)和会后出版物(会议结束后,经主办单位或其他单位整理、编辑出版的正式出版物)。

会议文献的主要特点是:传递信息比较及时,传递的信息针对性较强;它反映了某学科、专

业的最新成果和发展现状及趋势,是研究工作不可缺少的情报源。

④ 专利文献。专利是用法律来保护科学技术发明创造的制度。专利文献是专利制度的产物,一切与专利制度有关的各种专利文件统称为专利文献,其中包括发明说明书、专利说明书、专利局公报、专利文摘、专利分类与检索工具书,申请专利时提交的各种文件(如请求书、权利要求书、有关证书等)与专利有关的法律文件和诉讼资料等。狭义的专利文献一般是指专利局颁布出版的各种发明说明书或专利说明书及其所派生的各种二次文献。

专利文献的特点:数量巨大,覆盖面广;格式统一,措词严谨;描述对象具体、单一;技术内容新颖、可靠;文件类型多,重复量大,是重要的技术经济信息源。

⑤ 科技报告。科技报告是研究或设计单位向提供经费的上级部门提供的关于某项研究或设计任务完成情况及财务消耗情况的总结报告。

科技报告的特点:从形式上看,科技报告的出版形式比较特殊,每份报告自成一册,篇幅长短不等,有连续编号,装订简单,出版发行不规则。从内容上看,科技报告的内容比较新颖、详尽、专深。

⑥ 学位论文。学位论文是高等学校或研究机构培养的学生为获得某种学位而撰写的科学论文,一般有学士论文、硕士论文和博士论文。学位论文中除了少数可能发表在期刊或其他出版物以外,多数是不出版的。每篇学位论文有一复本保存在授予学位的学校的图书馆里,可供查阅。

⑦ 技术档案。技术档案是指在生产建设和科技部门的技术活动中形成的、有一定的工程对象的技术文件的总称。其内容包括:任务书、协议书、技术经济指标和审批文件;研究计划、方案、大纲和技术措施;有关的技术调查材料(原始记录、分析报告)、设计计算、试验项目、方案、数据和报告;设计图纸、工艺卡片以及应入档文件。

⑧ 产品资料。产品资料是指国内外各厂商为推销产品而印发的商业宣传品,包括产品样本、产品目录、产品说明书、厂商介绍、厂刊或外贸刊物、技术座谈资料等。

⑨ 标准文献。标准文献是以文件形式出现的、经过公认的权威当局批准的标准化工作成果,是可以采用文件形式或规定基本单位(物理常数)这两种形式固化下来的文件。标准化是为了有关各方的利益,特别是为了达到最佳的经济效果,并适当考虑到使用条件和安全要求,在有关各方的协作下,进行有步骤的特定活动所制定并实施各项规则的过程。

标准文献的特点:制定、审批有一定的程序;适用范围非常明确专一;编排格式、叙述方法严谨划一,措词准确;技术上具有较充分的可靠性和现实性;对有关各方有约束性,在一定条件下具有某种法律效率;有一定的有效时间,需要随着技术发展而不断修订、补充或废除。

⑩ 政府出版物。政府出版物是各国政府部门及其所属机构所发表的文件。它的内容广泛,概括起来可分为行政性文件和科技文献两大类。行政文件包括国会记录、司法资料、方针政策、规章制度、决议、指示以及调查统计资料等。科技文献包括各部门的研究报告、技术政策文件等。

(2) 按文献加工层次划分,可将科技文献分为零次文献、一次文献、二次文献和三次文献。

① 零次文献:即未以公开形式进入社会流通或未经出版发行的最原始的文献,如私人笔记、手稿、试验记录、会议记录、论文草稿、原始统计数据、考察记录等。零次文献的特点是,内容新颖,但不成熟,不定型,不公开交流,难以获得。

② 一次文献:即以作者本人的生产与科研工作成果为依据而撰写的、并已公开发行进入社会流通使用的原始文献,如专著、学术论文、科技报告、会议论文、专利文献、学位论文等。一

次文献的特点是,具有学术上的新观点、新发明、新技术、新成果,提供了新的知识信息;是创造性劳动的结晶,有直接参考、借鉴和使用的价值;是人们检索和利用的主要目标。

③ 二次文献:即将大量、分散、无序的一次文献收集起来,按照一定的方法进行整理、浓缩和加工,使之系统化而形成的各种目录、索引和文摘,是一种具有多种检索途径的检索工具。二次文献仅是对一次文献进行系统化的压缩,无新的知识产生,具有汇集性、检索性的特点。它的重要性在于提供了检索一次文献的线索,因此二次文献又称为检索性文献。

④ 三次文献:即根据一定的目的和需求,在大量利用一次、二次文献的基础上,对有关知识进行综合、分析、提炼、重组而再生的信息资源,如词典、手册、百科全书、年鉴、各种教科书及综述等。因此,三次文献具有综合性高、针对性强、系统性好、知识信息面广的特点。三次文献又称为参考性文献,有较高的使用价值,可直接参考、借鉴和利用。三次文献源于一次文献,又高于一次文献,是一种再创性文献。

从文献情报的角度看,零次文献是最原始文献的素材,一次文献是人们检索与利用的主要对象,二次文献是文献信息的检索工具,三次文献是人们考查数据、事实信息的主要信息源。

(3)按文献载体形式划分,可将科技文献分为印刷型文献、缩微型文献、机读型文献和声像型文献。

(4)按相对利用率划分,可以根据相对利用率的多少,把一个学科有关的文献分成核心文献、相关文献和边缘文献三类。核心文献通常是指与本学科发展水平、发展动向密切相关的一些文献。相关文献和边缘文献通常是指所含内容与本学科的关系相对疏远一些的文献。从学科发展的背景来说,相关文献和边缘文献是本学科与其他学科技术相互渗透、相互结合的结果。

三、电子信息源

1. 电子信息资源的起源和发展

电子信息资源是信息资源的一种,是电子化了的的信息资源,即以电子数据的形式,把文字、图形、图像、声音等多种形式的信息存放在光、磁等非印刷型介质上,以电信号、光信号的形式传输,并通过相应的计算机和其他外部设备再现出来的一种信息资源。

(1)起源——机编文献目录数据库信息阶段。

电子信息资源的真正起源是美国国立医学图书馆在20世纪60年代中期的电子计算机版的MEDLARS数据库,它被用于医学文献的检索。随后,美国其他一些机构也相继使用电子计算机编辑文摘期刊,从而客观上编制成了机器可读的书目文档,称之为文献目录型数据库。这种文献目录型数据库由于使用计算机来编辑、检索文献信息,因此可以将机编文献目录型数据库的研制视为电子信息资源形成的开端。

(2)国际联机检索信息阶段。

国际联机检索就是用户使用终端设备,远距离地从国际联机检索中心迅速而准确地获取电子文献信息,使知识信息得到广泛而有效地传播和利用,其实质是数据库和通信的结合。从20世纪六七十年代起,许多国家还先后建立了专门从事计算机检索的机构,如美国的洛克希德公司和系统发展公司,英国的目录检索服务处和图书馆自动化情报服务处与Infoline公司以及欧洲的DIMDI和ESA公司等,这些机构都建有大量的数据库联机检索系统,都向全世界联机用户提供电子信息服务。其中,著名的系统有Dialog、ORBIT、ESA-IRS、OCLC和STN等。