

现代信息 检索技术

主 编◎祖 彬 于纯良

XIANDAI XINXI JIANSUO JISHU

哈尔滨工程大学出版社

现代信息检索技术

主 编 祖 彬 于纯良

副主编 王 平 王丽华 马 晶 夏新兴

哈尔滨工程大学出版社

内 容 简 介

本书是编者根据多年的教学实践经验,结合当前信息科学的技术发展,汇集多方面有关资料编写而成。全书共分七章,内容包括文献基础知识、文献信息检索、文献检索工具、专利文献检索、计算机信息检索、数字图书馆电子图书数据库检索、INTERNET 信息资源及其检索。该书引证了近年来最新检索工具的体例和用法,并配有大量检索实例,内容系统,取材精练,适用性强。

本书读者对象为高等学校研究生、本科生和高职学生,可供从事高等学校教学、科研工作的同志学习、参考,也可以广泛用于信息中心、图书馆和情报咨询部门。

图书在版编目(CIP)数据

现代信息检索技术/祖彬,于纯良主编.—哈尔滨:哈尔滨工程大学出版社,2007.3

ISBN 978-7-81073-959-7

I. 现… II. ①祖…②于… III. 情报检索 IV. G252.7

中国版本图书馆 CIP 数据核字(2007)第 026341 号

出版发行 哈尔滨工程大学出版社
社 址 哈尔滨市南岗区东大直街 124 号
邮政编码 150001
发行电话 0451-82519328
传 真 0451-82519699
经 销 新华书店
印 刷 哈尔滨工业大学印刷厂
开 本 787mm×1 092mm 1/16
印 张 11.75
字 数 291 千字
版 次 2007 年 3 月第 1 版
印 次 2007 年 3 月第 1 次印刷
印 数 1—1 000 册
定 价 19.80 元
<http://press.hrbeu.edu.cn>
E-mail: heupress@hrbeu.edu.cn

前 言

21 世纪是信息技术快速发展的时代,整个世界在激烈的竞争中不断前进,信息化的程度已经被当作衡量一个国家综合国力的重要尺度,也是衡量一个国家的核心要素。在人类进入日新月异的信息社会的今天,信息的获取和利用成为人类赖以生存的本领,信息技能教育益发重要和迫切。

教育部曾于 1984 年就作出决定,要求各高等院校以图书馆为基地开设“文献检索与利用”课程,以适应科学技术发展,增强大学生的信息意识,培养和提高大学生利用文献和信息的技能。由此可见,加强文献检索课的教学,使学生了解信息源,快速地从浩如烟海的文献中获取知识和信息,不断增强自学和科研能力,已成为培养跨世纪人才的强烈要求。这为文献检索课的深入普及与发展带来了新的机遇,同时给文献检索课的教学也带来了新的挑战。为了适应高等职业教育改革的要求,突出职业教育的特色,满足高等职业院校对实用教材的需要,结合高职教学特点,编写了这本教材。

本书从现代信息检索的基本知识为切入点,全面系统地介绍了手工科技信息探索和计算机科技信息检索,并根据信息检索的发展,全面阐述了网络环境下文献检索的原理及其应用。全书共分为 7 章,内容包括:文献基础知识、文献信息检索、文献检索工具、专利文献检索、计算机信息检索、数字图书馆电子图书数据库检索、INTERNET 信息资源及其检索。全书内容系统,取材精练。本书在编写过程中,参考、借鉴了大量的文献检索教材及其他相关文献,也得到了许多同仁的关心和支持,在此一并致以诚挚地谢意。

本书由齐齐哈尔大学祖彬老师、烟台大学图书馆的于纯良老师担任主编。烟台大学的于纯良老师编写了第一章、第六章;齐齐哈尔大学的王平老师编写了第二章;齐齐哈尔大学的祖彬老师、王丽华老师编写了第三章、第四章;陕西科技大学的夏新兴博士编写了第五章;齐齐哈尔市劳动保护教育中心的马晶老师编写了第七章。全书由烟台大学图书馆的于纯良老师负责进行审阅。

由于编者水平有限,本书难免存在不妥之处,敬请各位同仁、专家和读者提出批评指正。愿本书能真正成为读者的良师益友。

编者

2007 年 1 月

目 录

第1章 文献基础知识概述	1
1.1 信息、知识和文献的概述	1
1.2 文献的种类	4
1.3 科技文献	7
1.4 学习科技文献检索的意义	10
第2章 文献信息检索	14
2.1 文献检索语言	14
2.2 文献检索工具	26
2.3 科技文献检索方法	28
2.4 科技论文的撰写	32
第3章 文献检索工具	44
3.1 文献检索工具的定义与基本职能	44
3.2 图书索引	44
3.3 期刊索引	47
3.4 中文网络数据库检索	50
3.5 外文网络数据库检索	72
第4章 专利文献检索	91
4.1 专利基本知识	91
4.2 专利文献	96
4.3 专利文献检索	109
第5章 计算机信息检索基础	112
5.1 计算机检索概述	112
5.2 信息存储与数据库结构	116
5.3 检索语言与检索技术	120
5.4 检索的基本程序	125
第6章 数字图书馆、电子图书数据库及其检索	128
6.1 数字图书馆概述	128
6.2 国外数字图书馆的发展现状	128
6.3 国内数字图书馆建设的发展现状	131
6.4 数字图书馆的资源建设模式	132
6.5 我国数字图书馆建设需要解决的主要问题	133
6.6 书生之家数据库	133
6.7 超星电子图书数据库	136
6.8 方正 Apabi 教学参考书数据库及其检索	146

第 7 章 INTERNET 信息资源及其检索	155
7.1 Internet 信息资源特点和信息资源种类	155
7.2 Internet 信息组织	157
7.3 Internet 信息资源检索	159
7.4 Internet 信息检索工具——搜索引擎	161
7.5 搜索引擎的种类	163
7.6 搜索引擎的一般查询规则	164
7.7 主要的中文搜索引擎	166
7.8 主要的英文搜索引擎	175
参考文献	182

第1章 文献基础知识概述

随着人类跨入 21 世纪,人类社会也由工业经济时代跨入了知识经济时代。知识经济时代的重要特征之一就是信息成为了时代的象征,这就如同材料、能源是工业经济时代的重要象征一样。在知识经济时代,信息作为一种重要的资源,已成为社会生产力的重要因素之一,科学、技术、经济以及一切的竞争都将以它的竞争为前奏,它对整个社会所产生的影响已初见端倪。一方面,由于知识经济时代的到来,信息变得越来越重要;另一方面,信息又以爆炸般的速度快速增长,变得越来越丰富,使我们如同置身于信息与知识的海洋当中。面对这样一个充满信息的世界,新知识不断涌现、知识新陈代谢频繁的世界,如何有效地、快速地掌握最新信息,获得最新知识,让信息为我所用,成为当前人们面临的重要问题之一。越来越多的人想从大量的文献中获取有用信息,但由于不懂检索技能,面对大量的垃圾信息陷入无奈的尴尬境地。这就告诉我们,面对未来的世界,如果没有一定的信息检索和利用的能力,就会面对汪洋大海般的信息陷入找不到、读不完的困境,从而无法获取最新信息和及时进行自我知识的更新,最终被社会所淘汰。因此,获取信息的能力对 21 世纪的高科技人才来讲必定会和外语交际能力、计算机应用能力一样成为衡量其基本素质的重要因素。

1.1 信息、知识和文献的概述

1.1.1 信息、知识和文献的概念

1. 信息的概念

信息是由事物发出的、体现它存在和运动状态的信号和消息。信息普遍存在于自然界、人类社会和思维方式中。

广义的信息是物质、能量存在和运动状态的表征。狭义的信息是人类社会共享的一切知识、学问以及从客观现象中提炼出来的各种消息的总和,即现代信息。

人类认识世界的过程,实际上就是不断地从外界取得信息和加工信息的过程,而人类改造世界的过程,是把加工外部信息所得的“主观”信息(如政策、计划等)反作用于外部世界的过程。

2. 知识的概念

知识是人类在改造客观世界实践中所获得的认识和经验的总和,是信息的一部分。人类认识的成果是在实践的基础上产生并经过实践检验的对客观实际的反映,由现象到本质,从感性到理性,从而掌握客观事物的规律性,产生知识。

3. 文献的概念

文献是记录有知识的一切载体,即用文字、图形、符号或声频等技术手段记录知识信息的物质载体,或称固化在载体上的知识信息。

1.1.2 信息、知识和文献之间的相互关系

信息、知识、文献之间的关系是事物发生发出信息;信息经人脑加工变为知识;知识被记录形成文献;文献经过传递、应用于理论与实践再产生新的信息。

1.1.3 信息检索概述

1. 信息的来源

一切信息来源于自然界,来源于人类社会。人们的生产、生活、科研以及政治活动等都是信息产生的源泉。

信息来源按其形式可分为口头(语言)信息源、实物信息源和文献信息源。

口头(语言)信息,在情报学领域称之为“零次信息”。

实物信息是以某种实际存在的物品所揭示出的信息,如样机、样品等。

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有目的地去搜集、加工、分析和利用。

文献信息是获取信息的主要来源。通过对文献上所记录信息的检索与利用,可以帮助我们打开知识信息宝库的大门。

2. 信息检索的概念

信息检索的全称是信息存储与检索,因此,广义信息检索概念包括两个方面:一是将信息按一定的方式组织存储起来,二是根据信息用户的需要找出有关的信息。对信息用户而言,后者更为重要。只有经过组织的有序信息集合才能提供检索。信息检索的本质是信息用户的需求和信息集合的比较与选择,即匹配的过程。

从用户需求出发,信息检索是对一定的信息集合采用一定的技术手段,根据一定的线索与准则找出相关信息的过程。

3. 信息检索的类型

人类在社会实践中,根据工作、学习、科学研究的不同需要对所进行的信息检索,大体上可归纳为三类:文献检索、数据检索和事实检索。

(1)文献检索主要是通过文献检索工具进行检索。例如:要查找超大规模集成电路研究信息,有哪些文献可以参考?查找这些文献的过程,就是文献检索。

(2)数据检索是通过参考工具书进行检索。例如:要查找某种电压调节器的参数,这就是数据检索。

(3)事实检索也是通过参考工具书进行检索。例如:要了解美国有哪些研究机构还在进行超大规模集成电路的研究,这就是事实检索。

信息检索流程图如下:

信息检索 { 数据、事实检索 → 参考工具 → 直接答案
 文献检索 → 文献检索工具(目录、题录、文摘、索引) → 间接答案

4. 文献信息源的特点

(1) 文献数量大,增长速度快

在17世纪以前有文字记载的数千年中,人类知识的积累非常缓慢。大约1000年翻一番。17~19世纪,由于印刷术的推广与应用,人类知识的记载量突增,大约每100年翻一番。自19世纪末以来,随着科学技术的突飞猛进,人类的科学知识量正以前所未有的速度急剧

增加。据推测,人类的科学知识在19世纪每50年增加二倍,20世纪中叶每10年增加一倍,20世纪70年代每5年增加一倍,80年代后每3年增加一倍。一些尖端领域和新兴学科,如计算机科学、核能、环境科学、分子生物学等,每2~3年文献量就要翻一番。

随着人类知识量的增加,文献的数量也正在以前所未有的速度迅速增加,世界各国的科学技术出版物在种类、数量、出版速度、出版形式等方面飞速发展。有关资料统计表明,科技成果每增长一倍,文献量要增长几倍;工业生产量每翻一番,文献量要翻四番。目前,全世界每年出版的图书达几万种,平均不到一分钟就有一本新书问世;全世界出版的科技期刊6万多种,每年发表科技论文约500万篇。平均每天发表含新知识的论文已达1.3~1.4万篇;全世界每年出版专利说明书(包括等同专利)达100多万件,累计已达3000万件;全世界的技术标准总数为20万件;每年出版的会议文献达1万多种,科技会议论文10万篇以上;每年公开出版科技报告20多万篇,学位论文20多万篇。除此之外,技术总结、调研报告、产品说明书、产品设计图纸、检测资料等不作发表的科技文献更是多得难以统计。

文献数量的激增,一方面表明文献信息资源的丰富,另一方面也给人们有效地选择、利用文献,获取所需信息造成了一定的障碍。

(2) 文献分布集中又分散

现代科学技术的日益综合与细化,使各学科之间的严格界限日趋淡化,学科之间的相互联系、交叉渗透逐渐增强。这使文献的分布呈现出既集中又分散的现象,即某一专业的大部分文章发表在少量的专业性期刊中,还有一部分则刊登在大量的相关专业,甚至不相关专业的杂志中。据对美国《化学文摘》进行统计发现:500种期刊集中了全部论文的62%;而全部论文的90%,分散在3000余种期刊中;余下10%的论文,则分散于另外的9000多种期刊中。文献的这种分布现象是普遍存在的,它提示人们在收集文献信息时,应首先选择本专业的核心期刊。

(3) 文献时效性增强

科技的迅速发展,使得新知识、新技术、新产品等层出不穷。这种现象加速了知识的新陈代谢,随之造成了文献的新陈代谢,使文献老化加速。有人统计各类文献的平均寿命为:图书10~20年,科技报告10年,学位论文5~7年,期刊3~5年,标准文献5年,专利文献短于法律规定的15~20年。

通常用文献的“半衰期”来描述文献的老化状况。所谓文献的半衰期是指某学科领域目前尚在使用的全部文献中,较新的一半所出版的年限。国外有人统计不同学科文献的半衰期为:地理学16.1年、地质学11.8年、数学10.5年、植物学10年、化学8.1年、生理学7.2年、机械工程5.2年、社会科学5年、化工科学4.8年、物理学4.6年、冶金学3.9年、生物医学3年。由于各国科技发展水平不同,相应的文献寿命也不相同。

当然,随着社会的进步,文献的增长量很大,内容会越来越丰富,但目前科技论文出现的总体水平下降,文章冗长、空洞无物、多次重复,是一种需要引起重视的不良倾向。

(4) 文献内容交叉重复

现代科技综合交叉、彼此渗透的特点,导致知识的产生和文献的内容也相互交叉、彼此重复。具体表现为:

同一内容的文献以不同文字发表。据统计,当前世界上每年翻译图书约占图书出版总量的10%;一些重要的核心期刊被译成多种文字在不同国家出版;同一项发明可以向多个国家申请专利,使专利说明书的内容出现严重重复现象,据世界知识产权组织统计,世界各

国每年公布的专利说明书的重复率高达 65% ~ 70%。

同一内容的文献以不同的形式出版。据调查,会议论文有 40% 经过会后整理、修改发表在期刊上;有相当数量的科技报告后来发表在期刊上,如美国的 NASA 报告有 80%、AD 报告有 60%、美国农业部的科技报告有 80% 既出版单行本又发表在期刊上;随着新型载体文献的普及应用,许多文献既出版印刷型的,又有缩微版、电子版等。

在激烈的商业竞争中许多畅销书内容雷同,选题重复。此外再版、改版文献数量不断增多,如 1992 年我国出版图书 92 148 种,其中新书仅占 63%。科技文献的交叉重复,可以给读者带来阅读的方便,但也造成了检索工作的困难。

(5) 文献载体及语种增多

随着声、光、电、磁等技术和新材料的广泛应用,新型文献载体不断涌现。传统的纸张型文献已失去了一统天下的局面,多种载体文献相互依存、相互补充、共同发展已成为趋势。新型的非纸张型文献如缩微型、机读型、视听型等,或增大了信息存储密度、延长了保存时间;或加快了信息传递与检索速度,实现了资源共享,文献家族日益繁荣。

各国文献所用语种不断增多。现在全世界出版的期刊所用语种达 70 ~ 80 种。据估计,目前全世界出版的科技文献有一半是用非英语发表的。

1.2 文献的种类

我国根据出版形式的不同,将文献分为如下几个类型。

1.2.1 期刊 (Periodical、Journal、Magazine)

国际标准化组织 (ISO) 给期刊的定义是: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。因此它是一种连续出版物。由于期刊的内容丰富而复杂,故又称为“杂志”。科技期刊的出版周期一般有周刊、半月刊、月刊、双月刊、季刊、半年刊等。科技期刊按内容可以化分为综合性的和专业性的。总的说来,其出版周期比较短,刊载文献的速度比较快,数量多,内容新,能及时反映某一学科领域的学术发展水平,所以很受科技人员的重视,成为科技人员首选的文献资料。目前,全世界共有 2 000 多个文种,用于发表文献的文种达 100 多种,其中核心期刊约 6 000 余种。

1.2.2 图书 (Book)

据不完全统计,我国科技图书出版量大约是每年达四五千种。大多数科技图书是著者在对已经取得的科技成果、成熟的生产技术知识和经验进行选择、鉴别、核对、组织而成的文字比较多而丰富的劳动成果。其论述比较系统、全面而可靠,阅读很方便。可分为阅读型和参考型,如各种教材 (textbook)、专著 (monograph)、丛书 (series of monographs)、手册 (handbook)、全书 (encyclopedia) 等,对于从事某一项专题的系统全面地研究具有较大的参考价值。主要问题是出版周期较长,因而难以满足人们对最新信息的需求。科技图书一般应属于三级文献,但也有一些专著含有部分新观点、新方法、新技术、新材料等,可能还具有一级文献的作用。

1.2.3 会议文献(Proceedings、Conference Papers)

会议文献在科技文献中具有特殊的地位,已经引起大多数科技工作者的重视,因为许多学科中的重要发现,有一大部分是在科技会议上作为首次公布的成果。

会议文献一般是指在学术会议上宣读的论文或书面发言,按照会议的等级可分为国际会议、全国性会议和基层会议。会议文献还可以分为会前文献、会中文献和会后文献三种。会前文献又可以分为:①会议情报性文献,用来预报将要召开的学术会议内容,主要有会议名称、涉及学科内容、地点、日期、发起机构名称和地址、截稿日期及会议出版物情况预报;②会前发表的论文预印本和论文摘要等,发给与会者阅读,一般为内部发行。会中文献主要有开幕词与闭幕词、会议决议书等。会后文献是在会议结束后,由主办单位整理编辑出版发行的会议论文集或会议丛刊。重要而有影响的在会后一般公开出版,主要形式有以下几种:①在某一固定期刊上刊载,或以特辑、增刊、专辑形式发表;②出版专题论文集(symposium),把论文汇编成册,取一个专门的书名,以图书形式出版,大部分会议文献以此种形式出版;③出版连续性会议文献(conference serial),有的直接按会议的届次排序,以会议主题名称出版丛刊(proceeding series)。

会前的会议文献主要有《世界会议:美国和加拿大》(World Meeting, United States and Canada)和《世界会议:美国和加拿大以外各国》(World Meeting, Outside USA and Canada),可以预报两年内将要召开的世界上学术会议的基本情况,包括会议名称、地点、召开日期、发起机构名称和地址、会议内容简介、截稿日期及会议出版物情况等。

会后文献检索工具有《会议论文索引》(Conference Paper Index)和《科学技术会议录索引》(Index to Scientific and Technical Proceedings)。

国内会议文献的检索工具有《中国学术会议文献通报》,它是检索我国召开的国内外学术会议资料的检索刊物。

1.2.4 专利文献(Patent Document)

专利文献是指在专利形成过程中产生的一系列官方文件和有关出版物的总称。广义专利文献是指专利局出版的与专利有关的各种文献。如专利公报、专利文摘、分类表、检索工具、说明书以及与专利有关的法律文献等。狭义专利文献仅指专利说明书,专利检索的主要对象是专利说明书。专利文献在形式上具有统一的格式;在文字上它是一种法律文件,文字较简练,特别是要求保护权利范围;在内容上具有广泛性、详尽性、实用性、新颖性、独创性以及具有较强的系统性、完整性和报道的及时性等特点。主要依据著录项中的专利号、专利国别、专利权人、专利优先日期(公开日期)、出版时间等加以识别。

目前全世界大约有 150 多个国家设立专利机构,70 多个国家出版专利资料。

1.2.5 科技报告(Technical Report)

一般是根据研究工作的进展情况、实验记录、最终结果写成的、能提供某一方面的完整技术。科技报告的类型主要有技术报告(Technical Report)、札记(Notes)、论文(Papers)、备忘录(Memorandum)、通报(Bulletin)等。

科技报告的特点是一个报告自成一册,有机构名称,有统一编号。科技报告出现于 20 世纪初叶,但自二战以来,由于国外不少军事、科研、工业机构利用政府研究报告对内、对外

传递科技情报,目前已经发展成为科技文献的一大门类。

严格说,科技报告都是一次文献;内容比较专、深,大致可以分为基础理论和生产技术两大类型。由于它是研究的记录和成果,代表了一个国家和某一专业的科技水平,因而可以对科研工作起到直接的借鉴作用。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。

1.2.6 学位论文(Dissertation、Thesis)

根据国际标准化组织(ISO)的定义,学位论文是指作者为取得专业资格而撰写的介绍本人研究成果的文献。学位论文在美国称为“Dissertation”,在英国称为“Thesis”。学位论文一般分为学士(bachelor)论文、硕士(master)论文、博士(doctor)论文等。1984年教育部(原国家教委)决定,我国所有研究生的博士和硕士论文,一律交中国科技情报研究所收藏,并提供中国学位论文的缩微平片。科技文献出版社发行的《中国学位论文通报》可以作为检索学位论文的工具。

1.2.7 标准文献(Standard Document)

标准文献一般是指各类技术标准、技术规范和技术法规等,包括国际标准(ISO或IEC等)、国家标准、专业标准、地方标准和企业标准等。标准文献是人们在从事科学研究、工程设计、生产建设、技术转让、国际贸易、商品检验中对工农业产品和工程建设质量、规格及其检验方法等方面所作的技术规定,是从事生产、建设和行政、组织管理时需共同遵守的具有法律约束性的技术依据和技术文件。

国际上比较著名权威的技术标准是国际标准化组织的ISO标准、国际电工委员会的IEC标准等。美国的ASTM标准是由美国材料与试验协会(American Society for Testing and Materials,简称ASTM)制定。该协会成立于1898年,是美国历史最悠久、规模最大的学术团体之一。ASTM出版物有50多年(1899~1950)累计索引可查,之后每隔5年又出版一本补充索引。它每年出版合订本,即《ASTM标准年鉴》,备有年度索引,便于检索使用。

此外,英国标准BS(British Standards),原联邦德国工业标准DIN(Deutsche Industrie Normen),日本工业标准JIS(Japanese Industrial Standards),澳大利亚标AS(Australia Standards)每年都有年鉴出版物。若工作中需要查询时,可以直接检索。

1.2.8 产品资料(Product Literature)

产品资料一般是指国内外厂商为推销产品而印发的商业宣传品,代表已投产产品,成熟可靠。大体有产品目录、产品说明书、产品数据手册等类型。这是对定性产品的性能、构造原理、用途、使用方法和操作规程、产品规格等所作的具体说明。

产品样本具有其他技术所没有的特色,从产品样本中,可以获得关于产品结构的详细说明。并且由于它代表已经投产的产品,在技术上比较成熟,数据上比较可靠,有较多的外观照片和结构图,直观性强。产品样本对于新产品的选型和设计都有一定的参考和借鉴作用,并可以从中学了解到世界生产动态和发展趋势,为我国进口机器设备或技术创新提供参考。需要注意的是,由于产品的不断更新,产品样本比较容易过时而失去先进性。

1.2.9 政府出版物(Government Publications)

政府出版物是各国政府部门及所属专门机构发表的有关文件。包括基础科学、应用科学直到政治、经济、贸易等社会科学。从文献性质看,可分为行政性文件(国会记录、政府法令、方针政策、规章制度、决议、指示、调查统计资料等)和科技文献(科研报告、科普资料、技术政策等)两大类。

1.2.10 其他(Others)

除上述出版类型外,其他文献还包括技术档案、工作札记、广播、电视、报纸等大众传媒、卫星资料,其中技术档案是指在生产建设和科研部门技术工作中形成的,有一定的工程对象的技术文件。如任务书、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、设计计算试验基础上的方案数据和报告等,具有重要的信息价值,有明显的保密性和内部控制使用的特点。

1.3 科技文献

“科技文献”就是“记录有科学技术信息或知识的一切载体”。

科技文献汇集着世代千百从事科学技术活动的人们辛勤劳动的成果,积累了无数有用的事实、数据、理论、定义、定律、定理、技术方法以及科学的构思和假想,记载了许多成功的经验和失败的教训,是人们从事科学研究和生产实验的历史记录,是劳动人民和科技工作者对客观事物认识的结晶。它反映当时人们对客观事物认识的程度和科学技术的进展状况及发展水平,预示着科学技术发展的趋势和方向,随着科学技术的产生而产生,并随之发展而发展。

1.3.1 科技文献的作用

首先,科技文献的第一个作用表现为:科技文献是科技进步的阶梯。这是因为科技文献记载了一代又一代人的劳动成果,保存了人类的精神财富,为后人进一步的科学研究提供基础。人们正是通过从科技文献中不断地吸取营养,批判地继承前人经验,扩大眼界,开阔思路,在已取得成果的基础上提出新问题,进而得出新结论,而攀上新的科学技术高峰。

其次,科技文献作为记录科技信息和知识的物质形式,其第二个作用表现在传递科技信息和知识。各个国家、各个部门正是通过科技文献的交流,使科技信息或知识得以广泛传播和充分利用,这充分体现了科技文献的继承性和国际性。

最后,科学技术的不断发展,使科技文献的数量不断增加,质量不断提高。反过来,科技文献的发展又加速了科学技术的进步,促进了社会的发展。因此,科技文献还有另外一个作用,即可以将科技文献的数量和质量作为衡量科学技术发展水平和成就的标志之一。

1.3.2 科技文献的分类方法

1.按信息载体划分(参考《GB 3469—83 文献类型与文献载体代码》)

(1)印刷本(printed form)

印刷型文献是以纸质材料为载体,以印刷为记录手段而形成的文献形式,是目前整个文献的主体,也是有着悠久历史的传统文献形式。它的特点是不需要特殊设备,可以随身携带,随处随时阅读。但存贮密度小,体积大,占据空间大,笨重,不利于保存。

(2)缩微制品(microform)

缩微型文献是以感光材料为载体,以照相为记录手段而形成的一种文献形式,包括缩微胶卷、缩微平片、缩微卡片等。缩微型文献的优点是体积小,便于收藏和保存,价格便宜等,但阅读时要有较复杂的阅读设备来支持。目前在整个文献中,所占数量较少,在一般的图书馆入藏亦较少。

(3)录音制品(audio form)和录像制品(visual form)

也简称声像型文献,它是以磁性和感光材料为介质记录声音、图像等信息的一种文献形式。包括唱片、录音磁带、电影片、幻灯片、录像磁带等。其优点是方便快捷,可闻其声,见其形,易理解。

(4)电子数字型(electronicform)

电子数字型文献是以计算机处理技术为核心记录信息的一种文献形式,包括计算机用磁带、磁盘、磁卡等。例如 120 mm 光盘可储存 540 MB 信息,但这类机读磁性制品需要配套设备。这种文献存储容量大检索速度快捷、灵活,使用方便。随着计算机技术特别是网络技术的迅猛发展和普及,电子数字型文献的地位越来越受到人们的重视。

(5)其他(others)

实物型展品,仪器、设备、文物等。

2.按级别划分

依据文献传递知识、信息的质和量的不同以及加工层次的不同,人们将文献分为四个等级,分别称为零次文献、一次文献、二次文献和三次文献。

(1)一次文献(primary document/literature)

这是人们直接以自己的生产、科研、社会活动等实践经验为依据生产出来的文献,也被称为原始文献(或叫一级文献),其所记载的知识、信息比较新颖、具体、详尽。一次文献在整个文献中是数量最大、种类最多、所包括的新鲜内容最多、使用最广、影响最大的文献,如期刊论文、专利文献、科技报告、会议录、学位论文等等,这些文献具有创新性、实用性和学术性等明显特征,是科技查新工作中进行文献对比分析的主要依据。

(2)二次文献(secondary document/literature)

二次文献也称二级文献,它是将大量分散、零乱、无序的一次文献进行整理、浓缩、提炼,并按照一定的逻辑顺序和科学体系加以编排存储,使之系统化,以便于检索利用。其主要类型有目录、题录、文摘、索引四类,如《中文科技资料目录》、《中国科技期刊数据库》等。二次文献具有明显的汇集性,系统性和可检索性,它汇集的不是一次文献本身,而是某个特定范围的一次文献线索。它的重要性在于使查找一次文献所花费的时间大大减少,二次文献是查新工作中检索文献所利用的主要工具。

(3)三次文献(tertiary document/literature)

三次文献也称三级文献,是选用大量有关的文献,经过综合、分析、研究而编写出来的文献。它通常是围绕某个专题,利用二次文献检索搜集大量相关文献,对其内容进行深度加工而成。属于这类文献的有综述、评论、评述、进展、动态等,这些对现有成果加以评论、综述并预测其发展趋势的文献,具有较高的实用价值。在查新工作中,可以充分利用反映某一领域

研究动态的综述类文献,在短时间内了解其研究历史、发展动态、水平等,以便能更准确地掌握待查项目的技术背景,把握查新点。

(4) 零次文献(zeroth document/literature)

这是一种特殊形式的情报信息源,主要包括两个方面的内容:一是形成一次文献以前的知识信息,即未经记录,未形成文字材料,是人们的“出你之口,入我之耳”的口头交谈,是直接作用于人的感觉器官的非文献型情报信息;二是未公开于社会,即未经正式发表的原始的文献,或没正式出版的各种书刊资料,如书信、手稿、记录、笔记和包括一些内部使用通过公开正式的订购途径所不能获得的书刊资料。

零次文献一般是通过口头交谈、参观展览、参加报告会等途径获取,不仅在内容上有一定的价值,而且能弥补一般公开文献从信息的客观形成到公开传播之间费时甚多的弊病。

3. 按文献获取方式的明朗程度划分

西方情报学者将文献分为三种类型:一种是“白色文献”,即通过书店、邮局等正常渠道可获得的文献,它是已知信息的载体;第二种是“黑色文献”,即根本得不到的绝密性资料,它是未知信息的载体;第三种便是“灰色文献”,顾名思义,灰色文献不容易得到,但又不是根本得不到,既含有已知信息,又有未知信息,加之文献具有非正式出版物的特点,在西方情报界又常被称为“难得文献”(Hard-to-get-literature)、“非常规文献”(non-convention literature)和“非正式出版文献”(non-publication literature)。

高等学校常用的灰色文献类型主要有:①内部刊物;②会议内部资料;③内部技术报告;④未出版的学位论文;⑤产品说明书、产品目录等有关资料。

灰色文献的特点主要有:①出版形式简单;②内容专一具体;③分散琐碎;④具有半封闭性。

4. 各文献级别间的关系

从零次文献、一次文献、二次文献到三次文献,是一个由分散到集中,由无序到有序,由博而精的对知识信息进行不同层次的加工过程。它们所含信息的质和量是不同的,对于改善人们的知识结构所起到的作用也不同。零次和一次文献是最基本的信息源,是文献信息检索和利用的主要对象;二次文献是一次文献的集中提炼和有序化,它是文献信息检索的工具;三次文献是把分散的零次文献、一次文献、二次文献,按照专题或知识的门类进行综合分析加工而成的成果,是高度浓缩的文献信息,它既是文献信息检索和利用的对象,又可作为检索文献信息的工具。

查阅文献则是通过二次文献检索到一次文献的过程。因此,二次文献是最重要的检索手段和工具,一次文献是最终的检索对象。

1.3.3 科技文献交流方式

1. 正式过程

借助于科技文献进行科技交流的过程。

2. 非正式过程

是指情报交流基本上是由科技人员自己来完成的那些过程,如信件、演讲、参观、交换资料 and 著作等的交流和讨论。它的优点是迅速性、高度选择性和针对性,反馈及时。缺点是适用范围有限、缺乏有效的社会监督(未经科技界评价的情报缺乏可靠性)、不便于加工整理、存储和积累。

1.4 学习科技文献检索的意义

我国教育部门早在 1984 年 2 月就印发了《关于在高等学校开设“文献检索与利用”课的意见》,又于 1985 年 9 月印发了《关于改进和发展文献课教学的几点意见》。这两个文件都明确指出:文献课不仅有助于当前教学质量的提高,而且是教育面向未来的一个组成部分。在高校中开设“文献信息检索”课,有目的地培养学生的信息意识和信息吸收能力,使其具备自我知识更新和获取最新信息的能力。

1.4.1 科技文献在国民经济建设中的作用

1. 科学技术是第一生产力

目前,世界物质财富大约以每 10 年按照 20% ~ 25% 的速度增加。科技文献大体是每 5 年要翻一番。有一些科技领域如信息科学,估计知识的半衰期只有 3 年。科技的迅速发展使产业结构、产品结构、生产工艺和劳动内涵很快调整、变化。如电信事业,今天的电报已经不多,代之而起的是传真(Fax)、电子邮件(E-mail)、移动电话、可视电话等。另外,像核电技术、杂交水稻和视听设备等技术的应用使人类在物质生活和精神生活方面有了一个飞跃。

科学技术是第一生产力,从科学技术飞速发展可以预示到,人们的工作将不会太固定,工作专业化的深入与综合化的扩展,科技与进步、繁荣与灾害、弊端并存,竞争与合作共生,以及国际化与地域化相结合,使得科学技术在国民经济生产中占据的地位越来越重要。

当然,科学技术本身并不等于现实的生产力,还需要转化成现实生产力。这一转化的媒介主要靠科技情报工作,靠科技文献。充分利用科技文献的功能,就会促进科学技术和经济的迅猛发展。

2. 在经济增长中占的地位

在过去经济不发达时期,多数的劳动工具是原始落后的,生产得到的产品主要靠的是劳动力(人)本身能力的大小。到了 20 世纪初,一部分工业化国家在经济增长诸因素中科技进步可以达到 5% ~ 20%,而在 20 世纪的 70 年代,科技进步的因素已经可以达到 60% ~ 80%。在目前的大多数发达国家,科技进步已经制约国家经济的增长,成为最重要的因素之一。

第二次世界大战后,日本的迅速崛起,显示出科技进步的作用。日本仅用了 15 年的时间就走完了西方先进工业国家半个世纪的历程,使其迅速跻身于世界经济强国之列。在市场经济条件下,科技、信息情报对经济发展起着至关重要的作用。任何个人、企业乃至国家,要想在竞争中站住脚,都必须掌握先进的科学技术和足够可靠的情报,并运用它进行科学决策。例如,日本的丰田汽车曾经一度要挤垮美国的第二大汽车公司,并不是日本的技术比美国高,日本人也承认他们的技术并不比美国强,而是在于情报工作。日本人发现美国造的汽车宽大、舒适、富丽堂皇,但是却费油。他们通过搜集情报,分析 10 年内必然出现石油的短缺,因此,查找搜集了世界各国有关汽车省油的资料,集中生产了小而经济省油的汽车,果然到了 20 世纪 70 年代末,日本的丰田汽车风靡世界,显示出优势。“以情报求生存,以情报求发展”已经成为我国当今众多有远见卓识的企业家的共识。

3. 由科技到生产物化周期缩短

由科学家实验得到的新技术、新产品到转化为人类广泛使用的物品,这个过程的时间随着科学技术的飞速发展而缩短了时间。

在18世纪,科学家经过研究认为,当时的物化周期在80~90年之间。到了19世纪的初中叶,物化周期约为30年左右,如电解冶炼铝的技术,从设想到实验成功,再到工业化生产用了35年的时间,这一技术使得铝制品被广泛应用。在20世纪的50年代,现代科学家们普遍认为物化周期平均在6~10年间,而有些技术则可能一年一个台阶,像计算机技术、通信技术可以说是日新月异。如电话用了56年,无线电技术用了35年,电视用了12年,大规模集成电路用了2年,激光技术仅用了1年,计算机技术基本上1~2年就上一个新的台阶。因此,利用文献知识武装头脑,跟上时代的步伐是非常重要的。

1.4.2 科技与文献的关系

经典力学的伟大的奠基人牛顿(Sir Isaac Newton, 1642~1727)临终前说过:“如果说我比别人站得高一些的话,那是因为我是站在巨人肩膀上的缘故。”牛顿就是在意大利的伽里略(Galileo Galilei, 1564~1642)和德国的开普勒(Johannes Kepler, 1571~1630)所认识、总结的力学定律的基础上,继承、借鉴、综合和发展了他们的学说,达到了经典力学的高峰。因此,可以说,科学技术与科技文献之间是相互促进、相互制约的密切关系。科技水平愈高,分工愈细,对文献需求愈迫切。

另一方面,情报工作愈好,可以更有力促进科技发展,因此,充分利用文献是打开科技成功之门的钥匙。像世界上登录的化合物1990年为1000多万种,到了2000年已经达到2000多万种,只有进行情报检索才能有的放矢地搞好科研工作。

据WIPO报道,在应用技术研究中,查文献可缩短研究时间的60%,节省费用40%。

统计材料表明,科研成果增一倍或产量翻一番,情报量增4~6倍。

1.4.3 学习科技文献检索的意义

1. 文献检索的重要性

(1) 节省查找文献的时间

据抽查,日本科技人员花在查找文献上的时间占整个科研时间的40%~50%,若改用计算机,在日本只需几分钟就可以从几百万件资料中找出有关资料。

(2) 避免重复别人的劳动

科学研究最忌讳的是重复。科研人员在选题阶段就应该进行必要的文献检索,从中了解到该项目提出的原因、历史状况、目前的进展状况如何等等。只有这样,科技人员才可以借鉴别人的劳动成果,直接进入实质性的研究阶段,避免重复研究,提高工作效率。例如在美国阿波罗宇宙飞船登月计划中,有一项钛合金空舱压力试验,他们用了20个钛合金空舱充甲醇做试验,结果因出现穿孔而报废,经济损失高达150万美元。事后才知道只要事先查一查美国的《化学文摘》就可以完全避免这次损失。因为早在10年前的文献中,就已经发现了解决这个问题的办法,只要在甲醇中加2%的水就行了。

我国某研究所用了大约10年的时间研制成功“以镁代银”新工艺,满怀信心地去申请专利,可是美国某公司早在20世纪20年代末就已经获得了这个专利。这项专利的说明书就收藏在该研究所所在地的科技信息所。

我国1956年人工合成蛋白(牛胰岛素)在世界上首次研究成功。这是因为该课题的科