

信息检索与利用

吴高魁 伊雪峰 主编

吉林大学出版社



信息检索与利用

主 编:吴高魁 伊雪峰

吉林大学出版社

内 容 简 介

本书是根据国家教委高等教育司[1992]44号文件及辽宁省教育厅[2002]87号文件对高校《信息检索与利用》课的教学要求,针对我校教学的实际需要而编写的。它简要地阐述了文献检索的基本原理和基本方法,较详细地介绍了国内外相关检索工具及使用方法,并介绍了计算机检索的原理和方法。为了便于大学生充分利用高校图书馆,又增加了有关高校图书馆利用指南方面的内容。全书内容力求简明扼要,有较强的针对性及实用性,并有丰富的检索实例,便于自学。它主要用作《信息检索与利用》课的参考教材,是高校师生掌握文献查找方法的速成参考书,也是针对刚刚迈入大学的新生进行图书馆利用入门教育的必修教材。

图书在版编目(CIP)数据

信息检索与利用/吴高魁,伊雪峰主编. —长春:吉林大学出版社,2005.12
ISBN 7-5601-3374-6

I. 信... II. ①吴...②伊... 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 148929 号

信息检索与利用

吴高魁 伊雪峰 主编

责任编辑、责任校对:刘子贵

封面设计:李维伦

吉林大学出版社出版
(长春市明德路 421 号)

吉林大学出版社发行
长春市日升印业有限公司印刷

开本:787×1092mm 1/16
印张:12
字数:323 千字

2006 年 4 月第 1 版
2006 年 4 月第 1 次印刷
印数:1-3050 册

ISBN 7-5601-3374-6

定价:23.50 元

前 言

当今社会是一个急剧变化的、竞争激烈的社会。处处有机会,时时有风险。信息就是机会,信息就是成功,信息就是现代社会发展进步的关键变量。人们为了在竞争中求得生存,在变动中求得发展,都需要获得各种有用的信息,而文献资料正是信息的主要来源;另一方面,大学生的自学能力,独立研究能力和创造能力在很大程度上都离不开对信息和知识的收集、选择、判断、分析和综合利用,而最大量的信息和知识就储存在文献中。因此,要培养学生独立获得知识和进行科学研究的能力,就必须使其善于检索和利用文献。据专家统计,一个大学生在校学到的知识只是他一生中所用知识的10%,其余90%是靠毕业以后的工作中不断地接受继续教育而获得的。“餐人以鱼,其食一饷;授人以渔,其食一生”。《信息检索与利用》课就是教人学习查找与利用文献方法的课程,掌握了这种科学方法,就可以不断地吸取新知识,改善知识结构,提高自学能力和研究能力,发挥创造能力。信息检索与利用不仅仅是一种知识,还是一种技能,更是一种科学方法,它是每个大学生和研究人员必须掌握的一种基本功。

为实现《信息检索与利用》课特定的教学目的,本教材以培养大学生的信息意识和信息能力为前提来构思教材结构及相应的教学内容,并力求使之能反映科学性、系统性和实用性。同时,根据这门课程具有实践性强的特点,对检索工具除阐述了编排原则外,还着重对各种检索途径的步骤和方法做了较详细地介绍,对当前国内外流行的检索新技术及新方法做了系统介绍。另外,为了便于对大学生进行图书馆利用入门教育,使大学生充分利用图书馆,还增加了图书馆利用指南方面的内容。愿本书成为你成功道路上的一块铺路石。

参加本书编写工作的人员及分工如下:

吴高魁 第5章 第6章 附录1

伊雪峰 第1章 第2章 第4章 第14章

刘欣 第3章 第11章 第12章 第13章

金桂花 第8章 第9章 第10章

刘英煜 第7章 第14章

由于编者水平有限,教材中如有缺点和疏漏之处,恳请读者批评指正。

编者

2005年2月18日

目 录

第1章 文献检索概述

- 1.1 信息与信息资源 1
- 1.2 信息源与科技文献 3
- 1.3 信息爆炸与文献检索 6

第2章 文献检索基础知识

- 2.1 检索工具 10
- 2.2 索引与检索语言 14
- 2.3 文献检索策略 24

第3章 中文检索工具

- 3.1 《全国报刊索引》 31
- 3.2 《中国机械工程文摘》 31
- 3.3 《电子科技文摘》 33

第4章 美国《工程索引》(EI)

- 4.1 《工程索引》概况 34
- 4.2 《EI》编排结构 34
- 4.3 《EI》检索方法 37

第5章 英国《科学文摘》(SA)

- 5.1 《科学文摘》概况 39
- 5.2 《PA》,《EEA》,《CCA》的编排结构及著录格式 40
- 5.3 检索途径与实例 45

第6章 《金属文摘》(MA)

- 6.1 《金属文摘》概况 48
- 6.2 《MA》月刊文摘本的编排及著录 50
- 6.3 《MA》月刊索引本的编排及著录 52
- 6.4 《Alloys Index》与《Steel Alert》简介 54
- 6.5 检索注意事项与检索实例 55

第7章 美国《化学文摘》(CA)

- 7.1 《化学文摘》概况 59
- 7.2 《CA》的结构与内容 60
- 7.3 《CA》索引的查用方法 61
- 7.4 《CA》的检索方法 66

第8章 美国《科学引文索引》(SCI)

- 8.1 《科学引文索引》概况 68
- 8.2 《SCI》的结构及著录格式 68
- 8.3 《SCI》的检索方法与检索实例 72

第9章 美国《科学技术会议录索引》(ISTP)

- 9.1 《科学技术会议录索引》概况 74
- 9.2 《ISTP》的结构及著录格式 74
- 9.3 《ISTP》的检索方法与检索实例 77

第10章 美国《社会学文摘》(SA)

- 10.1 《社会学文摘》概况 80
- 10.2 《SA》的结构及著录格式 80
- 10.3 《SA》的检索方法与检索实例 83

第11章 中国专利文献检索

- 11.1 专利基础知识 86
- 11.2 专利文献的作用及特点 87
- 11.3 国际专利分类法 87
- 11.4 中国专利说明书的种类和编号体系 88
- 11.5 中国专利检索工具 89
- 11.6 中国专利检索方法 90

第12章 德温特专利检索工具

- 12.1 德温特出版物概况 92
- 12.2 WPIG 的各种索引格式 93
- 12.3 《文摘周报》的著录格式 96
- 12.4 德温特专利的检索方法 97

第13章 参考工具书

- 13.1 参考工具书的概念及特点 99
- 13.2 辞书 99
- 13.3 百科全书 103
- 13.4 年鉴与手册 107
- 13.5 名录与资料汇编 110
- 13.6 利用参考工具书应注意的事项 112

第14章 计算机检索

- 14.1 计算机检索概述 114
- 14.2 计算机信息检索基础知识 116
- 14.3 数据库检索系统 123
- 14.4 INTERNET 信息资源的检索 158
- 附录1 沈阳工业大学图书馆利用指南 168
- 附录2 黑本式拉丁字母-日文字母音译对照表 184
- 附录3 俄文字母-拉丁字母音译对照表 185
- 参考文献 186

第1章 文献检索概述

1.1 信息与信息资源

当今世界,信息已被视作现代社会的重要战略资源,信息资源的充分开发和有效利用已经成为社会经济发展的重要推动力。

信息资源不同于其他资源,如果说物质资源向人类提供的是材料,能量资源向人类提供的是动力,那么信息资源向人类提供的则是非物质形态的社会财富——宝贵的信息、情报和知识。信息资源的开发和利用,可以使有用的信息进入到国民经济和社会生活的各个领域,不仅可以节约物质资源和能量资源,而且还可以提高整个国民经济的素质,节约劳力、资金,扩大财富的增殖空间,提高产出效率,使国民经济系统创造出更多的社会财富,满足人们日益增长的物质和文化生活的需要。

1.1.1 信息及其属性

(1) 什么是信息

对于信息,人们并不陌生。无论在什么时候,也不管在什么地方,人们总是离不开信息,不是在收集信息、感知信息,就是在加工信息、处理信息。那么,究竟什么是信息呢?简单地说,信息就是事物的状态和特征。它普遍存在于自然界、人类社会以及人们认识和思维过程之中,人类生活的世界就是一个充满信息的世界。

信息现象是一种非常古老的现象。在人类历史的黎明到来之前,信息就已经存在于物质世界。凡是有物质、有能量的空间,就有信息。阳光普照、星斗灿烂,是宇宙天体发出的信息;电闪雷鸣、山呼海啸,是大自然发出的信息;物质的微观结构内部的各种射线,也是物质发出的信息;等等。这些自然界发出的信息叫作自然信息。在人类社会诞生之后,信息存在于自然界,也存在于人类社会;信息来自于物质领域,也来自于精神领域。人类认识和改造自然及客观世界的过程,就是一个信息过程。从人的认识和行为来看,人通过感觉器官,对事物的运动状态和方式进行感知,即信息接收,他所感知和了解的就是信息,然后将所获得的信息通过神经系统传输给大脑,即信息传递,大脑对信息进行分析、综合、加工判断,即信息处理,作出决策,再通过神经系统把决策传递给执行器官,最后由执行器官作出行动、反映,并将行动、反映的结果由感官再一次感知,反馈给大脑,即信息反馈。自有人类活动以来,就存在着这种信息现象。这些人际间传播的信息叫社会信息,社会信息是人类活动的产物,又是人类从事各种活动并期望达到预期目标的基础和前提。信息按其性质不仅有自然信息和社会信息之分,还有记录信息(文献信息)与非记录信息、语音信息与非语音信息之分。

人类虽然很早就接触和利用信息,但对信息的科学认识却是二十世纪以来的事情。信息不像物质那样的实物,对于什么是信息,人们有不同的说法。在我国,据《新词源》考证,一千多年前,唐代就曾有“梦断美人沉信息,目穿长路倚楼台”、“塞外音书无信息,道旁车马起尘埃”的诗句,其中“信息”一词的意思是音信、消息的意思。英语中的“信息”一词是“Information”,从词源上来分析,它是由“In”和“formation”两部分构成,前者是收到的意思,后者是整理成章的意思。这就是说,“信息”一词的意义是“把不明确的东西弄清楚之后,理顺成章,再传给需要的人”。所以,在英文词典中,通常把“知识”一词作为信息的一个功能同义词来使用。在我国的《辞海》

中,把信息解释为对消息接受者来说预先不知道的报道。日本出版的《现代用语基础知识》一书对信息的定义是:生活主体同外部客体之间有关情况的消息。美国传统词典把信息定义为:从学习、经验感受或教育中获得的知识。

(2)信息的主要属性

信息有许多重要属性。其中最基本的是知识性、可传递性和时效性。

①知识性。信息与知识有非常密切的关系,它是创造知识的材料,而且它本身有时也可以是一种知识。在一定的历史条件下,人们通过有区别、有选择的信息,对自然界、人类社会、思维方式和运动规律进行认识和掌握,并通过大脑的思维使信息有序化,形成知识。因此,从这个意义上说,信息是知识的原料,知识是有序化的信息。如果信息接受者接收的是经过加工处理的有序化信息,那么这种信息本身就是一种知识。人类社会的进步,就是人们根据获得的信息来感知世界、认识世界、改造世界的过程,也就是创造知识、利用知识、积累知识和发展知识的过程。因此,信息是重要资源。

②可传递性。信息总是依附于某种载体(如文献、物体、声音等),故能在一定的时空范围内传递。信息在时间上的传递称为存贮,在空间中的转移称为通信。信息的这个属性是非常重要的,它使人类的知识可以积累和传播,使人与人之间能够进行信息交流,使人与其环境之间能保持信息的联系,从而可以更好地认识世界和改造世界。

③时效性。信息的效用表现在它可以消除事物的不确定性。但信息具有动态性,即一切活的信息都是随时间而变化的,因此,信息是有时效的。脱离源体的信息,不再反映变化源体新的运动状态和方式,因此它的效用就会降低或完全失去效用。所以人获得信息之后,并不能满足,更不能一劳永逸,信息要及时发挥效用,知识要不断补充和更新。

1.1.2 信息资源

信息作为一种资源,从狭义的角度来理解,是指人类社会活动中经过有序化加工并大量积累后的有用信息的集合,从广义的角度来理解,信息资源还包括信息资源生产者和信息技术的集合。

信息资源是社会发所必须的一种重要的战略资源,为人类提供的是非物质形态的社会财富。世界新技术革命的发展表明,工业革命已让位于信息革命,世界经济正朝着依靠更多的信息资源投入,生产出物质消耗更少,质量更好,更加耐用的产品的经济结构方向转变,信息资源与物质资源、能量资源一并共同构成现代社会经济发展的三大支柱,信息资源的开发与利用已经成为社会经济发展的重要推动力。

在激烈的国际竞争中,信息资源已经成为人们争夺的重点,谁能更多更快地占有信息资源并能有效地开发和充分利用,谁就能作出正确的决策,取得国际竞争的优势,创造出经济腾飞的奇迹。二十一世纪将出现划分强国与弱国的新标准,将出现“信息强国”和“信息弱国”之分。某些小国和弱国将受控于某些信息强国,将沦为这些信息大国的“信息保护国”或“信息殖民地”。全球将出现“信息争夺战”。目前发达国家竞相开展的“信息高速公路”建设已经拉开了“信息争夺战”的序幕。信息已经成为生产、竞争力和经济成就的关键因素,它是一种战略资源。尽管它不是唯一的资源,但却是最重要的资源。

信息资源不同于其他资源的最大特征是它可以为人们所共享。信息可以脱离事物而独立存在,也可以荷载于其他载体,被无限地进行复制,传播和分配,为大家共享和利用。信息不仅可以被多次使用、多次开发,而且不会在使用中消耗掉。物资资源和能量资源不具备这种特征,它们不能被人们所共享。

我国信息资源经过五十余年的建设已经取得了显著的成就。据一些科研报告和论文提供的数据,仅仅是我国400多个独立的科技信息机构,以及国家专利、标准等部门馆藏的各类科技期刊、图书,科技报告、会议录、专利、标准、样本等,累积已达数亿万件以上。如果考虑各类图书馆、大学、科研机构的馆藏,这一数量还要大得多。我国检索刊物体系基本形成,到目前已有近140种检索期刊,年报道量近100万条。自80年代初,电子化信息资源有了较快的增长,引进的国外文献磁带52种,资源总量达到2000万条,引进的文献光盘数量也超过100种。国际联机终端保持了一定数量分布在50个城市,约有100台。我国目前拥有自建数据库系统超过1000个,信息资源总量数十亿条。特别是二十世纪90年代初以来,随着信息网络的发展,信息资源开始上网,而且这一趋势正在加快。此外,信息资源合理布局与共享得到了推进,信息资源管理水平及开发利用水平不断提高。但是,这些成就与工业发达国家相比,仍有较大的差距,主要表现在信息资源总量不足,布局分散,开发利用信息资源的水平不高。

邓小平同志曾经指出“开发信息资源,服务四化建设”。所以在新世纪之初,加强信息资源的建设、开发利用,正是为四个现代化创造环境和条件,是对国民经济信息化最积极的响应和参与。

1.2 信息源与科技文献

1.2.1 信息源

人们通常把产生信息的源泉或母体简称为信息源。从事科学研究工作或生产管理及经营活动的组织和个人是一种信息源,各种图书馆,信息中心是一种信息源,文献资料本身也是一种信息源。从根本上来说,一切信息都来源于自然界或人类的实践活动,只有自然界和人类社会才是信息的真正源泉。文献上记录的信息只能称为信息流,是信息传递过程中的一种存在方式或表达方式。不过,对于大多数信息用户或读者来说,文献资料确是他们获取信息的主要来源。所以习惯上也把它们统称为信息源。

为了便于存贮和传递,信息总是要依附于某种载体。所以,信息源按载体形式可分为:口头信息源、文献信息源和实物信息源。

口头信息源——它以人的声音为载体,信息提供者或发送者直接用口头谈话方式将信息传出去,如各种报告会、新闻发布会或个别交谈等。口头信息源具有传递速度快、选择性强、反馈迅速等优点。其缺点是直接传播面较窄,信息易逸失,稍纵即逝,而且难以实行有效的社会监督。

实物信息源——它附着某种实际的物品(如产品、样品、样机等)上面。信息用户通过参观或考察来采集其中的有用信息。其优点是直观、真实,易检验或仿制。但一般需经过复杂的分析或解析过程才能将其中的有用信息分离出来,各种展览会获取实物信息的重要渠道。

文献信息源——信息以文字、符号、图形等形式记录在纸张、感光材料或磁性材料上,以出版印刷的方式传递出去。这是一种最大最重要的信息源。信息用户通过阅读、理解来吸取其中的有用信息。它传播面广,便于系统积累和长期保存,是在时间和空间上积累和传递知识的最有效手段,也是人们获取信息的主要来源。

1.2.2 科技文献

所谓科技文献,是指用文字、符号、图形、声频、视频等技术手段记录科技信息或知识的载体。它是重要的信息源之一。

科技文献按不同的划分标准,有不同的类型。

(1)按科技文献外在的物质形态划分,科技文献可分为:

①印刷型:指传统纸张印刷品。印刷方法有铅印、油印、胶印等。纸张印刷品的优点是,便于流传阅读,不受时间、地点和条件的限制,缺点是比较笨重,不易保存。主要品种有图书、期刊、报纸等。

②缩微型:是指通过现代技术手段将所需文献缩小复制到胶片等材料上,使原来的文献缩小几十倍甚至上千倍。它的特点是,体积小,重量轻,占有空间少,便于保存和自动化检索等。其缺点是不能直接阅读,需借助于缩微阅读机。

③机读型:是指计算机可读型文献,主要有磁带、磁盘等。它是将文献内容变成计算机可读的语言储存在磁带或磁盘上,阅读时由计算机输出变成人们能读懂的语言。

④声像型:包括唱片、录音带、录像带、光盘、幻灯片、科技影片等。这些资料声情并茂,给人以直观感受。声像文献在某些方面能起到独特的作用。如语音教学等。

(2)按图书情报人员对文献的加工深度划分,科技文献可分为:

①一次文献:是以著者本人在科研、生产工作中取得的科技成果为依据,撰写、创作的原始文献,不论其载体形式、出版类型如何,都属于一次文献。例如,期刊论文、科技报告、专利说明书、会议文献等,都是一次文献。一次文献直接记载了科研和生产中创造发明成果的原始资料,是科技人员宝贵的参考资料。

②二次文献:一次文献类型多,数量大,不便于检索利用。将分散的,无组织的各类一次文献收集起来,进行加工、整理、简化和组织,形成文献的目录、索引和文摘,作为检索一次文献的工具,这样的检索工具即为二次文献。利用二次文献可以帮助科技人员全面、系统地查找有关一次文献的线索,节省查找文献所花费的时间,尽快弄清楚某一课题的发展历史、现状和趋势,避免造成重复劳动,准确了解与课题有关的技术资料,借鉴别人的经验和教训,加快科研和生产的步伐。

③三次文献:是在利用二次文献检得一次文献线索的基础上,合理应用一次文献,对其内容进行综合、分析、评述,再度加工的产物。例如:图书专著、评论、综述、手册、年鉴、百科全书等,都属于三次文献。三次文献一般综合性强,系统性好,知识面广,有的还具有检索功能,参考价值比较大。

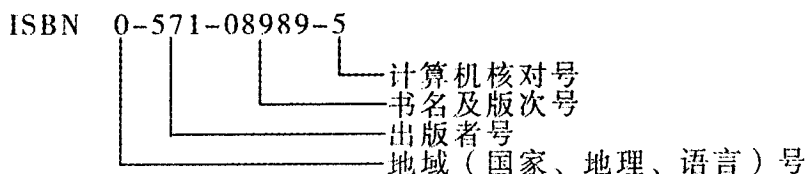
以上三者之间的关系是:一次文献是科技文献的基础,是文献检索的主要对象,二次文献是一次文献的简化,是检索科技文献的工具,三次文献是一次文献的高度浓缩,是重要的信息源。

(3)按出版形式划分,科技文献有十大类型:

①科技图书:科技图书是对科学技术成果、生产技术知识和经验的概括和总结。包括科技著作和参考工具书。科技图书提供的资料,内容比较成熟、系统,能全面系统地了解某一特定领域中的历史和现状,可以将人们正确地领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。

现在许多正式出版发行的图书在版权页上印有国际标准书号 ISBN(International Standard Book Number),它是国际标准化组织 ISO 制订的图书国际编号。ISBN 具有唯一性,指的是一种图书,如果装帧不同,版本不同,就有不同的 ISBN。下面举例说明 ISBN 的含义。

ISBN 由 10 位数字组成,例如:0-571-08989-5,具体含义解释如下:



要想获得某一册科技图书,必须正确注明书名、著者、出版社、出版年月、国别、文种及版本等项目。

②科技期刊:期刊也叫杂志,是指采用固定名称的定期或不定期出版的连续性刊物。每期有连续的卷、期号或年、月顺序号,有固定的篇幅和开本,内容一般最少有两篇以上的独立文章。科技期刊出版量大,学科多,内容新,是及时了解科技动态,掌握某一学科的进展情况,开阔视野,拓宽思路,吸收新知识的好工具。

许多期刊的封面上印有国际标准刊号 ISSN(International Standard Series Number)。ISSN 也有唯一性,指的是一种期刊,不管哪一种期刊,ISSN 始终不变。例如:ISSN 1000-5420,其中前七位是标准号,最后一位是计算机核对号。要获得某一期刊,必须正确注明该期刊的刊名、卷、期、出版年月、出版国家或出版单位及文种。

③科技报告:科技报告又称研究报告、技术报告,是作为科研工作成果的正式报告,或者是对研究和试验过程中的每个阶段的进展情况的实际记录。这种报告在形式上,每份报告自成一册;在内容上,题目专深具体,往往涉及尖端学科的最新研究课题;在时间上,一般都早于期刊等其他类型的文献。科技报告是在第二次世界大战期间及战后迅速发展起来的,并逐渐成为一种重要的信息源。目前全世界每年出版的报告大约 60 万件以上,其中公开发表的约有 20 万件,其余的都是保密报告。美国有著名的 PB、AD、NASA 及 DOE 四大报告。报道内容各有侧重,PB 报告的内容侧重于民用工程技术方面;AD 报告侧重于军用工程技术方面,NASA 报告侧重于航空和宇航方面;DOE 报告侧重于能源及其应用方面。四大报告每种都包括数十万篇,占全世界科技报告的绝大多数。要索取科技报告,必须有报告号。

④会议文献:科技会议已成为科学技术交流的一条重要渠道。各个科学技术学会、协会、有关主管部门经常召开学术会议,会议文献就是这些会议的报告、记录、论文集等。这些文献代表着一门学科或专业的最新研究成果,反映了国内外发展水平和趋势,因此是信息的重要来源。会议文献,除专题论文集、连续会议录外,常常以期刊特号、专题或编入科技报告中发表。如果以会议录(图书)的形式报道,则需要掌握会议召开的时间、地点、会议名称、会议录名称,会议录出版年、出版地、出版单位,论文所在的页码;如果会议论文是以期刊的形式报道,则需要掌握该期刊的刊名、年、卷、期,以及刊登论文的页码;如果是以报告的形式报道,则需要掌握报告号。

⑤专利文献:专利文献是一种用法律形式来保护的文献。专利文献包括有:专利说明书、专利公报、专利文摘、与专利有关的法律文件及诉讼资料等。这是对专利文献的广义解释。狭义的专利文献就是专利说明书。专利说明书是专利文献的主体。专利说明书是指个人或机构为了获得某项发明的专利权,在申请专利时必须向专利局呈交的有关发明的一份详细技术说明书。它说明该项目发明的目的、用途、特点、效果及采用何种原理或方法等方面。

专利文献不仅向人们提供维护工业社会秩序的权利情报,而且向人们提供大量的技术开发中不可缺少的技术情报。世界各国的科技人员都十分重视专利文献,利用率非常高。一般在下列情况时需要查专利文献:查明某些技术领域内新发明和新设计的发展动向;制定长远的技术

开发计划;调查先进技术;从别人的发明或设计中得到启发;避免对研究和生产设备的重复投资;获得与外国公司技术合作的谈判资料,加强自己的谈判地位,等等。如果想获得一份专利文献,必须知道专利说明书的编号。

⑥标准文献:是指由技术标准及其他在特定活动领域内必须执行的规格、定额、规划、要求等的技术文件所组成的一种特定形式的技术文献体系。标准文献有一定的法律约束力。标准文献的种类有国际标准、区域标准、国家标准、部颁标准、专业标准、企业标准等。技术标准已成为各国执行技术政策所必需的工具。通过标准文献可以了解各国的经济政策、技术政策、生产水平、资源情况和标准化水平;先进的标准可供研制新产品、改造老产品,改进工艺和提高操作水平时借鉴;外贸方面的检验工作也需以标准文献为依据。要获得某一技术标准,必须正确注明该技术标准颁布的国家或标准化组织名称,制订年份及标准号。

⑦政府出版物:是指各国政府部门及专门机构发表、出版的文献。它的内容广泛,概括起来可分为行政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、决议等。科技方面的包括各部门的研究报告、技术政策文件等。其中科技文献占整个政府出版物 30% ~ 40%。尽管政府出版物在出版之前,部分已被所在单位出版过,所以与其他文献(如科技报告等)有时重复。但是政府出版物对于了解某国的科技政策,经济政策及演变情况,对于了解其科技活动本身,具有一定的作用。

⑧学位论文:学位论文是作者为了取得各级学位进行公开答辩而撰写的学术论文。在撰写论文过程中,既有导师的监督指导又有严格的科学规范及要求。学位论文的质量较高,信息价值较大。但此类论文属非卖品,不易搜集。

⑨产品资料:产品资料是各国厂商为推销产品而印发的宣传性出版物。其中包含有关产品的技术信息,如产品的性能、规格、重量、结构图、线路图及照片等。它是生产科研单位研究、分析各国产品技术发展情况和产品水平的重要资料。也是贸易部门判断某项产品价值或洽购产品的主要依据。产品资料图文并茂给人以直观印象,对造型设计、技术引进等都有较大参考价值。另外产品样本一般都作为赠阅品,而且传递迅速,交流广泛。

⑩其他类型:科技文献除了上述列出的几种主要类型外,还有许多类型。这里统归为其他类型。如报纸、科技档案、技术图纸、技术资料等。

1.3 信息爆炸与文献检索

1.3.1 信息爆炸

当我们着手去搜集所需要的信息时,马上就会感觉到现有的信息种类繁多,数量巨大,分布很分散,搜集难度很大。

据联合国教科文组织统计,全世界每年发表的科技论文约有 500 万篇,出版图书 50 万种以上,期刊 10 万种以上,专利文献 100 万件以上。另据中国图书进出口公司统计,全世界有大小出版社约 12 万家,每年出书约 60 万种,报刊约 18 万种(其中较有参考价值的约为 3 万种),各种文献资料约 400 万件以上。

最近几十年以来,科技文献的数量急剧增长。据有些资料估计,科技文献数量大约每十年翻一番,且其倍增周期有进一步缩短的趋势。文献使用的语言也在增多,目前已有几十种,比较通用的有 12 种。据英国情报学家霍森估计,目前全世界的科技文献有一半是用非英语发表的。另外,随着科学技术发展速度的加快,文献的有效期也愈来愈短,失效速度也逐渐加快。

人们称这种现象为“信息爆炸”或“情报爆炸”。它是由于科学发展、科学对新技术的推动

以及人们对信息和知识的需求日益增长等一系列因素相互作用下造成的。其后果是:人们对汹涌而至的信息洪流感到难以应付,许多有价值的信息还未来得及发现和利用就消失得无影无踪了。继之而来的是信息不灵,决策失误,研究重复,浪费巨大,搜集信息的时间消耗之大也使人感到难以忍受。

古人云:“书山有路勤为径。”而在信息爆炸的今天,要获取一切有用的知识,光靠一个“勤”字恐怕不够了,还要再加上一个“巧”字,即既要勤奋学习,又要掌握搜集信息和利用信息的技巧,其中包括文献检索或信息检索的方法。本书将重点介绍各种文献信息的检索方法,以期对大家的学习和研究工作有所帮助。

1.3.2 文献检索

(1) 什么是文献检索

为了便于确切地解释文献检索这个概念的含义,我们先解释一下什么叫信息检索及信息检索的类型。

信息检索来源于英语 Information Retrieval 的译义,是信息工作和文献工作的术语,是指将信息按照一定方式组织和存贮起来,并根据需要找出所需特定信息的过程。它的全称又叫信息存贮与检索(Information Storage and Retrieval),这是广义的信息检索。狭义的信息检索,是指从信息源中查找出所需特定信息的过程。据此,可以将信息检索表述为从庞大的高度分散的信息源中查找出所需特定信息的过程。

信息检索根据不同的检索对象,一般可分为三种类型:文献检索(Document Retrieval),数据检索(Data Retrieval)和事实检索(Fact Retrieval)。

文献检索:凡以文献(或文献线索、文摘)为检索对象的叫文献检索。检索结果是特定文献。例如,检索“关于通讯卫星都有哪些文献?”就是文献检索。

数据检索:凡以数据(或计算公式、化学分子式等)为检索对象的叫数据检索。检索结果是数值性数据。例如,检索“通讯卫星的飞行高度,飞行速度各是多少?”就是数据检索。

事实检索:凡以事实为检索对象的叫事实检索。检索的结果是某一事实,例如,检索“世界上有哪些国家发射了通讯卫星?”就是事实检索。

从信息检索的三种类型来看,文献检索与数据检索和事实检索之间在原理、方法和实践方面,并无本质差异,相反,有着密切的联系。数据和事实要通过检索的文献才能获得。因此,文献检索是最基本的形式,是获取信息的主要手段。正因为如此,国内外有不少人常把文献检索与信息检索等同起来,把文献检索作为信息检索的同义语。可见,文献检索只是信息检索的一部分,但它是其中最重要的一部分。

(2) 文献检索的重要性

文献检索的重要性主要体现在促进信息资源的开发和利用;协助管理者作出正确的决策;便于继承和借鉴前人的成果,避免重复研究或走弯路;节省研究人员查找文献的时间等方面。

① 促进信息资源的开发和利用

人类社会正跨入“信息时代”。信息是社会中的重要资源和宝贵财富。信息资源的开发利用已经显露出越来越大的经济效益和社会效益。正因为如此,二十世纪60年代以来,许多发达国家一直致力于推行“信息化”政策,通过开发利用信息资源来促进经济发展。进入90年代以来,发达国家信息产业的产值已占国民生产总值的40~60%,年增长率是传统产业的3~5倍。我国每生产1美元产品所消耗的能源、原材料为发达国家的2~5倍,而消耗的信息流量却只有世界平均水平的1/10。我国要在近几十年内建成有中国特色的中等发达的社会主义国家,也

必须重视信息资源的开发与利用。目前,我国政府已发出推进国家信息化的号召,并指出信息化的核心内容是开发利用信息资源。文献检索是开发利用信息资源的有力工具,如果把信息资源看作是一个巨大的宝库的话,那么文献检索就是打开这个宝库的一把“金钥匙”,它能帮助人们传播知识和利用知识,使知识转化为社会物质财富或创造出更多的精神财富,为促进社会的发展贡献力量。

②协助管理者作出正确的决策

信息虽不能确保决策正确无误,但它是决策的基础。《孙子兵法》云:“知己知彼,百战不殆。”打仗是如此,搞管理也是如此。在今天这个复杂的社会,管理问题显得尤其重要。从一个国家(地区或单位)来说,要发展什么,限制什么,引进什么,都需要有准确、可靠和及时的信息,以便作出正确的决策。日本引进大量国外先进技术,促使其经济“腾飞”,重要原因之一是技术信息搞得较好,及时准确地掌握了国外科技发展动向。而我国前些年由于信息不灵,盲目引进,使一些引进技术和设备不适用,造成了惊人的损失和浪费。遗憾的是,至今还有不少管理者没有认识到信息工作对决策的重要性。如果要花费气力、金钱或时间去获取信息,那么他们宁愿在无知的情况下采取行动。这样的决策和行动难免会造成一些不良的、甚至是灾难性的后果。例如,由于我国多年来一直未能解决的重复建设问题,至使现在很多产品出现了供过于求的情况,VCD的生产就是一个实例,生产的数量位于世界第一,造成了产品的大量积压。再如,我国贵州省的普安县,由于信息不灵造成的经济损失极大,仅1994年,普安一个个体种植户种植的生姜,因信息不灵,卖不出去,全部烂掉,造成直接经济损失10多万元;另外盲目上马的县酒厂,经济损失100多万元;兴办的大理石厂,损失10多万元,等等。可见,作为一个管理者,如果平时重视信息工作,善于利用文献检索等信息检索方法搜集信息,并对这些信息进行分析,就有可能做出正确的决策,避免重大的损失和浪费。

③便于继承和借鉴前人的成果,避免重复研究或走弯路

整个科学技术史表明:积累、继承和借鉴前人的研究成果是科技发展的重要前提,没有继承就不可能有创新。所以,牛顿说:“假如我比别人看得略为远些,那是因为我站在巨人的肩膀上。”

在研究工作中,任何一个项目从选题、试验研究或设计到成果鉴定,每一步都离不开情报信息。只有充分掌握了有关情报信息,知道哪些工作前人已经做了,哪些目前正在做,什么人在做,进展情况如何,这样才能避免重复,少走弯路,保证工作在高水平上进行,并缩短研究周期,获得预期的经济效益。相反,如果继承和借鉴工作做得不好,“闭门造车”,就容易造成重复研究;或者采用了较差的实验方法或程序,使研究工作进展缓慢甚至失败。

例如,国外一位科研人员,搞了一项利用计算机控制汽化器的装置A。用传感器监控内燃机速度、进气压、温度,由一数字运算器对传感器信号进行处理,达到自动控制进气的目的。申请专利时未被批准。因为日本日立公司已掌握这种技术,并已在美国申请专利得到批准。这项已取得专利的电子控制汽化器,精度和效率都比这位科研人员搞的装置A高。又如,国内某单位可关断可控硅(GTO)科研取得了一定的成果。但存在使用时容易烧毁晶体闸流管的问题。承担这项研究的科研人员不会查找国外有关文献资料,还以为是国外机密。试验长期在烧管换管中徘徊。后来他们与文献检索有经验的人员一起查找资料,找到一批对口文献。从这些文献中很受启发,总结经验,修改后再试验不再烧管子了。经过不断改进完善,有关指标达到国际先进水平,获得了市级一等科技成果奖。

④节省研究人员查找文献的时间。

据调查,科研人员为了确定科研课题,进行科学研究,以致最后成功,在全过程所花费的时间中,有相当比例是用于检索需要的文献信息。如美国曾对 8000 名化学化工科研人员调查,统计用于文献检索的时间比例,最长达 61%,最少是 15%,平均为 33.4%。日本某电气公司科研人员实验研究时间占 42%;用于检索、计划研究、资料处理的时间占 58%。从国内外的许多实践经验看,科研中出现的各种问题,几乎有 85 - 99% 需要而且可以通过科技文献检索获得启发、帮助和解决。因此,只有借助文献检索,才能在最短的时间内获得所需的信息,腾出更多的精力搞研究,提高科研效率。

思考题

1. 如何理解信息的涵义?
2. 简述信息的主要属性。
3. 科技文献的出版形式有哪些?
4. 信息检索有几种类型?

第2章 文献检索基础知识

当人类社会步入信息时代时,由于科学技术迅速发展,一方面科技文献的数量日益庞大;另一方面广大科技人员对信息需求日益迫切,因此,科技文献的庞大数量与科技人员对信息的迫切需求之间存在着很大矛盾。而解决这一矛盾的最好办法就是文献检索。为此就要学习和掌握必要的检索基础知识。这不仅可以帮助科技人员加深对检索的认识和理解,而且还可以用来指导其科学、准确地进行文献检索。文献检索基础知识涉及面很广,从应用角度出发,需要学习的内容有检索工具、索引与检索语言和检索策略等几个方面。

2.1 检索工具

检索工具是存贮、报道和查找文献的工具(或者说检索工具是以浓缩形式报道、存贮和查找信息线索的工具)。存贮文献不是存贮文献的全文,而是把有信息价值的文献的某些内容特征和外部特征著录下来,成为一条条文献线索,并将它们排列成有条不紊的整体。这个存贮文献的过程,也就是将一次文献加工成为二次文献的过程;另一方面,检索工具具有必要的检索功能,人们按照一定的方法,就能从检索工具中检出所需文献的线索。这个过程就是检索过程。存贮的目的是为了检索;要检索就必须存贮,二者互相依存。任何一部检索工具,都具有存贮和检索这两个方面的功能。

从检索角度看,检索工具是人们查找文献不可缺少的工具。如果把大量珍贵的科技文献比作一座知识宝库,那么,检索工具好比钥匙,掌握它,就能打开知识宝库的大门,获取到所需的科技文献。

2.1.1 检索工具的类型及特点

(1) 检索工具的类型

检索工具的类型很多,按照不同的分类标准可以划分成不同的类型。

①按照收录范围划分:有综合性、专业性和单一性检索工具。

②按照著录格式划分:有目录、题录、文摘等检索工具。

③按照出版形式划分:有卡片式、期刊式、单卷式、附录式、胶卷式、磁带式 and 光盘式等检索工具。

④按照检索手段划分:有手工检索工具和计算机检索系统。手工检索工具是指仅用手工方式来处理和查找文献的一切工具。如目录、题录和文摘等检索工具。计算机检索系统就是用计算机来处理和查找文献的一种电子化自动化的系统,由计算机、检索软件、数据库、检索终端及其它外围设备组成。

(2) 检索工具的特点

检索工具是用来存贮和检索文献信息的工作系统,因此它应具有三个必备要素:

①必须存贮有丰富的文献款目:存贮文献款目是指详细著录的文献题名,文献著者、文献出处和文摘等一条条记录。

②必须具备必要的检索标识:是指提供检索的各种标志,如分类号,主题词,著者姓名,文献编码等。

③必须具有提供多种途径的检索手段:是指提供检索途径的辅助手段。如主题索引、著者

索引等各种索引。

总之,作为检索工具,必须具备上述三个条件,否则就不能称其为检索工具。

2.1.2 检索工具的内容结构

检索工具都是由一定数量的文献著录款目,按照严密的结构编制而成的,典型的检索工具,一般由说明、正文、辅助索引和附录四个部分组成。

(1) 说明

说明是检索工具的必要组成部分,它包括封面、书名页、版权页、目次、前言或后记等。编者通过这些内容,向用户说明该检索工具的编制目的,内容范围、收录年限、著录说明、使用方法和对象等,为用户使用提供了必要的指导。因此,用户在利用每种检索工具时,必须首先仔细阅读检索工具的说明,熟悉和掌握其特点和使用方法,避免不必要的差错,提高检索的效果。

(2) 正文

正文是检索工具的主体部分,由所收录文献信息的著录款目组成。这是因为编制检索工具时,为了节省存贮空间,防止检索工具体积过大,因而编入检索工具的不是文献信息的全文,而是描述文献外表特征和内容特征的著录。每一篇文章的著录称为款目,每一条款目著录项目多少,根据检索工具编制方式不同而异。但是,不论什么检索工具,款目的著录项目至少应包括文献题名、著者和出处等几项主要内容。检索工具收录的文献越多,著录的款目也就越多,众多的文献款目汇集起来,按照一定的方式编成文献检索工具或文献数据库。信息检索就是从文献检索工具或文献数据库中查出相关的文献著录,从而根据文献的著录了解其内容和获得原文的线索。因此说,正文是检索工具结构中的主体。

检索工具正文部分的检索功能优劣,关键在于著录款目的编排是否科学合理。正文编排一般有分类、主题、篇名、著者等方式。前两种属内容编排,具有推荐性强的特点,适用于按一定专题查阅成批文献的要求。后两种属于形式排列,具有专指性强的特点,适用于查找特定文献的要求。检索人员能否顺利地从正文中获取所需文献线索,重要的是要根据检索目的和要求,区别检索工具的编排特点。

(3) 辅助索引

辅助索引是检索工具的重要组成部分,一部好的检索工具必须能提供迅速、准确、全面查找所需文献线索的多种途径。辅助索引就是为了扩大检索途径,加强文献检索深度而编制的。一般由主题索引、著者索引、文献编码索引组成。

辅助索引在检索工具中,虽处于辅助地位,但从在检索工具中发挥的作用来看,则处于重要地位,辅助索引种类越多,则检索途径就越多;辅助索引是体现检索功能的重要标志,能否打开检索工具存贮的文献宝库的关键是辅助索引。

(4) 附录

附录是检索工具的内容的补充,一般大型检索工具都有附录,主要包括收录文献类型、摘引文献目录、术语缩写、语种对照以及文献收藏单位等表。因此,附录为检索人员提供了必备参考资料,帮助检索人员进一步了解检索工具,识别文献专用术语,掌握文献来源。

2.1.3 目录性检索工具

目录是编排最简单、查找最方便的一种检索工具,它在信息检索中有着一定地位。因此,熟悉和了解目录的基本概念,掌握目录的基本类型及特点和作用,对文献检索具有重要意义。

(1) 目录及目录款目