

高等学校教材

供基础、临床、药学、预防、护理、精神类专业用

信息检索与利用

主 编 薛平安 任淑敏

常务副主编 孙思琴 楚存坤

副 主 编 明 红 张 宁 张念华

张保成 张清文 黄 皓

编委会成员 (以姓氏笔画为序)

王倩飞 田永晓 吕少妮 朱启贞 任淑敏

孙思琴 张 宁 张念华 张保成 张清文

明 红 黄 皓 楚存坤 薛平安 魏 涛

人 民 卫 生 出 版 社

图书在版编目 (CIP) 数据

信息检索与利用/薛平安等主编. —北京: 人民卫生出版社, 2008.9

ISBN 978-7-117-10546-0

I. 信… II. 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 (2008) 第 124532 号

信息检索与利用

主 编: 薛平安 任淑敏

出版发行: 人民卫生出版社(中继线 010-67616688)

地 址: 北京市丰台区方庄芳群园 3 区 3 号楼

邮 编: 100078

网 址: <http://www.pmph.com>

E-mail: pmph@pmph.com

购书热线: 010-67605754 010-65264830

印 刷: 潮河印业有限公司

经 销: 新华书店

开 本: 787×1092 1/16 印张: 16.5

字 数: 391 千字

版 次: 2008 年 9 月第 1 版 2008 年 9 月第 1 版第 1 次印刷

标准书号: ISBN 978-7-117-10546-0/R·10547

定 价: 40.00 元

版权所有, 侵权必究, 打击盗版举报电话: 010-87613394

(凡属印装质量问题请与本社销售部联系退换)

目 录

第 1 章 概论	1
1.1 信息、信息社会与信息素质	1
1.2 文献信息概况及其特点	4
1.3 信息检索及其重要性	8
第 2 章 信息检索基础	12
2.1 信息检索系统及检索语言	12
2.2 信息检索工具	16
2.3 信息检索步骤及检索效果评价	19
第 3 章 计算机信息检索概述	22
3.1 计算机信息检索原理及发展	22
3.2 Internet 概述	25
3.3 Internet 信息资源及检索工具简介	29
3.4 计算机检索基本技术与检索策略	36
第 4 章 中文题录、摘要型检索工具	42
4.1 全国报刊索引数据库	42
4.2 中国生物医学文献数据库	47
第 5 章 外文文摘型检索工具	58
5.1 MEDLINE 光盘数据库	58
5.2 PubMed 检索系统	77
5.3 美国《化学文摘》	88
5.4 Engineering Village	97
第 6 章 中文全文数据库	101
6.1 全文数据库检索基础	101
6.2 中国知识基础设施工程	102
6.3 万方数据	110
6.4 维普信息资源系统	117
第 7 章 外文全文数据库	125

7.1	OVID 检索系统	125
7.2	Springer Link	128
7.3	EBSCOhost	134
第 8 章	引文信息检索	141
8.1	美国《科学引文索引》	141
8.2	《中国科学引文索引》数据库	147
第 9 章	专利信息检索	155
9.1	专利和专利文献	155
9.2	中国专利信息检索	158
9.3	国外专利信息检索	162
第 10 章	循证医学信息检索	168
10.1	循证医学概况	168
10.2	循证医学证据及分级	169
10.3	循证医学信息的检索	170
第 11 章	互联网医学信息检索	176
11.1	医学搜索引擎及网站应用原则	176
11.2	医学专业搜索引擎	177
11.3	常用医学网站	182
11.4	互联网免费电子期刊查询	189
11.5	互联网上参考工具	193
11.6	人与机构查询	202
第 12 章	图书馆文献资源的利用	206
12.1	图书馆职能、类型及书刊查找	206
12.2	参考工具书的应用	207
第 13 章	科技查新	218
13.1	科技查新的产生与发展	218
13.2	科技查新概述	219
13.3	科技查新工作程序	222
13.4	查新质量评价	225
第 14 章	信息的利用	228
14.1	文献信息的获取	228
14.2	文献信息调研	232

第 15 章 科技论文写作基础.....	244
15.1 科技论文类型与特征.....	244
15.2 科技论文的选题与资料收集.....	247
15.3 科技论文的撰写与格式.....	250
附录 1 科学技术部科技查新咨询机构名单.....	254
附录 2 卫生部医药卫生科技项目查新咨询单位 (31 家)	255
附录 3 教育部部级科技查新工作站.....	256

第1章 概 论

社会飞速发展的今天,人类已经进入一个信息化社会和知识经济时代,在这个社会里知识和信息成为推动科技进步和社会发展的重要因素,对于任何人、任何组织和国家,获取和利用知识、信息的能力成为竞争成败的关键。但要从浩如烟海、瞬息万变的信息海洋中能够及时获取所需信息,必须掌握一定的知识和技能。本章内容从基础开始,主要介绍信息、知识、文献的概念、特征和相互关系以及获取利用知识、信息的意义及重要性。

1.1 信息、信息社会与信息素质

信息作为当今社会的三大资源(物质、能量和信息)之一,其价值性和重要性日显突出。因此,人们获取信息的能力成为生存、发展必不可少的组成部分,而信息素质强弱则影响着信息的获取和利用。

1.1.1 信息

信息(information),是事物存在方式、运动状态及其特征的反映,是事物发出的信号、消息。宇宙由万事万物组成,不同的事物有着不同的特征、存在方式和运动状态,因此也就会产生各种各样的不同信息。

信息普遍存在于自然界、人类社会以及人的思维活动中。例如,昼夜的变化是一种信息,它反映了地球绕太阳自转的运动特性和状态;树的年轮是一种信息,它反映了树木生长的时间特性;在医学上,病人各种症状、体征的出现和变化、各种检验的数据结果、X线图像显示都是疾病反应的信息。

信息是无形的,但它是客观存在的,是人们认识事物的基础。人们通过对事物产生的各种不同信息,进行分析综合,来认识事物的本质。人类发展的历史就是不断获取信息、传递信息、利用信息和创造信息的过程,就是通过信息来认识世界和改造世界的过程。

1. 信息的特征 信息主要具有以下四大基本特征。

(1) 共享性:信息源于事物,但不是事物本身,是人们认识事物的媒介。同时信息能够共享,即同一内容的信息在同一时间、同一地域可以被多个用户分享,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耗或消失,这是信息不同于物质和能量的最重要特征之一。

(2) 存储性和传递性:存储性和传递性是信息的基本特征之一,利用信息可存储性,人们可以有意识地将信息以某种方式存储在一定的物质载体上,使信息与物质载体构成

依附性很强、相对稳定的关系,避免信息的流失,成为人们可以检索、利用的信息源。信息的传递性是指信息能借助一定的传输工具和载体进行传递,形成信息联系,被人们感受和接受。正是由于信息的存储性和传递性,使人类能够掌握更多的经验和知识,推动了人类文明的进程。

(3) 价值性:信息作为一种资源与物质、能量共同构成支撑当代社会发展的三大支柱。物质资源为社会提供了丰富的原材料,能量资源提供了形式多样的动力,而信息资源则为社会提供了无穷无尽知识与智慧。它向人类提供生产知识的原料,是一种促进科技、经济和社会发展的新型资源。信息作为一种资源,被人们利用后能启迪思想,增进知识,改变知识结构,提高认识世界和改造世界的能力,并产生一定的社会效益和经济效益。

信息是一种资源,但并非所有的信息都能成为资源,只有那些经过人类开发与重新组织后的信息才能成为信息资源。信息资源是指信息世界中对人类有价值的那一部分信息,是附加了人类劳动,可供人类利用的信息。信息需要经过一定方式的传递,才能被人们接收和利用,信息只有被利用才会产生价值,否则其价值随时间流逝而减少,甚至成为“信息垃圾”。

(4) 时效性:人们获取信息的目的在于利用,信息的价值与信息利用时间密切相关,适时的信息才能产生最好的效用,使用迟缓就会降低效用或无用,甚至造成损害。信息描述的是事物,产生的是知识,事物在变化,知识在发展,随着时间的推移,原有的信息和知识与事物的状况就可能会出现某种程度的不符,从而逐步过时老化,失去其价值。信息的时效性决定着信息价值的大小,因此要求人们在搜集、加工、提供和利用信息时要及时。

2. 信息的来源 信息分为文献信息和非文献信息两大类。

文献信息是人类用文字、符号、数据、图像、声频、视频等方式记录在一定载体上的信息(以文献形式被记录的信息),它可以跨越时空无限循环地为人类所利用。

非文献信息包括口语信息、体语信息和实物信息三种。口语信息通常是以讲演、授课、讨论、广播等口头方式表述未被记录的信息;体语信息是在特定的文化背景下,以表情、手势、姿态等方式表述的信息;实物信息则是以样品、模型、雕塑等实物形式表述的信息。文献信息是主要的信息源,是人们获取、利用信息的重要渠道,因此也是本书讨论的重点。

1.1.2 知识

人类在认识和改造客观世界过程中,不断地发现和接受事物发出的信息,大量的信息经过人脑思维,进行分析、综合,获得了对事物本质和规律的认识,总结出了经验,也就产生了知识。因此可以认为:知识(knowledge),是优化、理论化了的信息集合。知识是对信息进行提炼和深化的结果,是人类对自然界、人类社会以及思维方式与运动规律的认识、辨析、理解和掌握,是人的主观世界对客观世界的概括和真实反映,是人类社会实践的总结。

知识来源于信息,信息是知识得以形成和传播的中介,信息不等于知识,知识是大量信息经过人的大脑加工处理后的产物。后人利用前人积累的知识来指导生产实践、科

学研究,在实践中又获得新的知识,知识的发展带来了人类生产、生活、环境从量到质的转变。由信息获得知识,同时又产生信息,这种更高形式的循环,使得信息愈来愈丰富,知识越来越全面,越来越深化,进一步提高了人们认识世界和改造世界的能力,从而不断推动社会向前发展。

21世纪已经进入以知识经济为主导的知识型信息化社会,与以往经济形态最大的不同点在于:它不是直接取决于有形资源、资本、硬件技术的数量、规模和增量,而是直接依赖于知识或有效信息的积累和利用,所取得的成果都是信息、知识改变资源的结果。知识的发展带来了人类生产、生活、环境从量到质的转变,知识本身就是财富。在这样的社会里,财富的积累、经济的增长、社会的进步、个人的发展都要以信息、知识为基础。知识在人类社会的发展中起着巨大的作用,是衡量一个民族、一个国家文明程度的标志。

由于划分标准不同,可以将知识划分成不同的类型,国际经合组织(OECD)将人类现有的知识分为四大类:

- (1) Know What (对象性知识,知道是什么),关于事实方面的知识;
- (2) Know Why (价值性知识,知道为什么),关于自然原理和规律方面的知识;
- (3) Know How (技术性知识,知道怎样做),关于技能或能力方面的知识;
- (4) Know Who (主体性知识,知道谁有知识),关于到哪里寻求知识的知识。

1.1.3 信息社会与信息素质

当今世界是一个发展迅速、竞争激烈、优胜劣汰的信息化社会,人们面临的是一种以信息为标识的全新的思维模式。同时二十一世纪又是一个高科技的社会,它需要人们有与科学技术发展相适应的完整的知识结构和技能。在这个社会里,知识和信息成为科技进步和社会发展的决定性因素,知识与信息的重要性比以往任何时候都明显。《第三次浪潮》的作者、未来学家托夫勒有句名言:“未来的世界只有两种人,一种是掌握信息的,一种是不掌握信息的,统治者与被统治者的区别是知识和信息量的区别。”因此人们提出了这样的口号“信息就是力量,信息就是机会,信息就是成功”。获取知识和信息的能力已成为人们生存、发展必不可少的基本技能。信息意识强的人,都把搜集和利用信息作为事业的生存之本、取胜之道。

生活在信息化社会的大学生,要提高综合素质,适应社会发展的需要。因此在综合素质教育中,信息素质作为其中一部分显得尤为重要,其素质的强弱不仅关系到大学生在校期间学习、掌握知识的多少,更重要的是关系到他们终生的业务能力和科研能力的提高。较高的信息素质,能有效地改善文化素养和思想方式,增强创新能力,及时调整知识结构,有效地适应信息社会的需要。

信息素质既是一种能力素质,更是一种基础素质。信息素质包括信息意识、信息能力和信息道德三方面内容。信息意识表现为个体对信息重要性认识程度和对信息是否具有特殊的、敏锐的感受力和持久的注意力;信息能力主要是对信息的搜集获取能力、分析鉴别能力和综合利用能力;信息道德是指个体在信息的获取、使用、制造、传播过程中应遵守的伦理规范。上述三方面内容中信息意识尤显重要,信息意识的强弱决定了获取、判断和利用信息的自觉程度,一个对于信息的重要作用有深刻认识的人,必定对信

息需求非常敏锐,敏锐的信息需求必定产生强烈的信息检索行为,对通过检索获得的信息也必定会加以认真地吸收和充分的利用。其次是信息能力,能力是知识的体现,知识是能力的基础,具备有比较全面的信息知识基础,才能具备一定的信息能力。信息道德则是强调的在获取和利用信息的活动中不得危害社会 and 他人合法权益的基本准则。

教育部要求开设的文献检索课,就是培养大学生信息素质的重要措施。通过文献检索课的学习和实践,增强信息意识,了解应遵守的信息道德,掌握常用的检索工具使用方法,提高学生收集处理信息的能力、获取新知识的能力和分析问题解决问题的能力,进而提高科研创新能力和社会竞争的能力。

1.2 文献信息概况及其特点

文献是最主要的信息源,检索文献是获取信息、知识的主要手段,而文献类型复杂,数量大、更新快,分布广泛又分散交叉,需要掌握文献的类型、特点,才能更好的获取所需信息和知识。

1.2.1 文献

“文献,记录有知识的一切载体”(中华人民共和国国家标准 GB3792.1—83)。具体地说,文献(literature, document),是将知识用文字、符号、图像、音频等记录在一定的物质载体上的结合体。

从这个定义可看出构成文献有四个基本要素:①知识,这是文献的核心要素,是文献的实质内容和灵魂所在,没有知识内容,文献也就不存在。②载体,负载知识的物质载体是文献的外在形式,因为知识、信息符号要落实在一个物质载体上,通过一定的方式和手段固定在一定的载体上才能显现出来,才能让人们看的见,摸得着。③文字、图像、符号、音频、视频,这是有形的知识、信息,只有通过它们才能将知识、信息表达清楚,才能记录在载体上。不通过它们,知识、信息就无法表达,也就不能称之为文献。④记录,是联系知识与载体的手段,是把知识、信息用文字、图像、符号等记载在一定载体上的动态过程,只有通过记录,才能把上述孤立的三要素结合起来成为文献,记录方式反映了科学技术发展和物质生产方式的水平。上述四要素共同组成了文献这个有机的、不可分割的整体,缺少其一都会失去文献的含义和特征,都不可能成为文献而失去保存、传递知识信息所特有的功能和价值。

文献是社会发展的产物。人类在漫长的生产实践、科学实践和社会实践中逐步认识了客观世界规律,从而产生了大量的知识,同时又运用已取得的知识再认识、再实践,不断地改造世界。为了能把在认识自然和改造自然过程中产生的知识积累、保存起来,便于传播和利用,人们就想办法把它固定在一定的载体上,以物化的形式保存下来,这就产生了文献。

文献是知识宝库和重要的信息源。文献是人类长期从事生产和科学技术活动以及社会交往的真实记录,它记录了人类历史长河中科学技术发展和人类社会活动的成就及水平,凝集着千千万万人的辛勤劳动,是人类智慧的结晶。文献中积累了许许多多有价值的事实、数据、理论、方法和科学的假象,记载着许许多多成功的经验或失败的教训,

反映着一个时代、一定的历史条件下科学技术的进展和水平,也预示着未来的发展趋势和动向。几千年来,文献为人类的文明进步做出了巨大的贡献,一直是人类了解过去、认识现在和预测未来的重要工具,对于人类社会的发展起着承前启后的桥梁作用。文献的出现和发展,突破了人类交流的空间和时间障碍,加速了知识的再生产和提高了知识的利用率,是人们获取知识、信息的重要来源。

文献是社会发展的产物,是人类文明的真实记录和反映,同时又是推动历史发展和社会进步的强大动力。

综上所述,我们可将信息、知识、文献三者之间的关系可归结为:信息是事物存在方式、运动状态及其特征的反映,是生产知识的原料;知识是被人类优化、理论化后的信息集合;文献是存储和传递信息、知识的载体,是重要的信息源。

1.2.2 文献的类型

根据不同的划分标准,文献分为不同类型。

1. 按文献载体划分 自商周以来,文献形态几经变迁,其载体形式由甲骨、金石、竹木、帛绢转而纸张,直至现在的胶片、磁带、磁盘、磁鼓、光盘等。记录手段也由人工刻、铸、书写到印刷再到今天的光、电、磁时代,文献及与之相关的各个方面发生了巨大变化。目前以载体划分的文献类型主要有:

(1) 手写型文献,是指人工书写或抄写而成的文献。如书写在竹签、帛或纸张上的古代文献、书法作品、书信、手稿、原始记录等。

(2) 印刷型文献,是指以纸张为载体通过胶印、铅印、石印、油印、影印等手段形成的纸质出版物。

(3) 缩微型文献,是以感光材料为载体,以照相为记录手段,把文献缩小形成的复制文献。有缩微平片、缩微胶卷等。

(4) 视听型文献,又称声像文献,是利用声像技术将声音、图像等记录在磁带、胶带、光盘、磁盘上,通过播放技术给人以听觉、视觉感受的文献。有音带、像带、幻灯片、电影拷贝等。

(5) 电子型文献,又称机读文献、数字文献,是把知识记录在计算机存储介质上的出版物,是通过电子计算机进行阅读的文献。由于电子文献的出现,加之网络通信技术的应用,使文献信息时效性更强、内容更广泛、检索也更加快捷、方便。

2. 按文献加工程度划分 依据文献传递知识、信息的质和量以及加工层次的不同,人们将文献分为以下几种。

(1) 零次文献:这是一种特殊形式的文献信息源。是指未公开出版的实验记录、原始录音(像)、书信、手稿,口头交流的信息或实物等被称之为“零次文献”。这一特殊的文献信息,具有内容新颖,零散性、不成熟性和难以检索的特点。

(2) 一次文献:这是作者直接以自己的工作、科研、社会活动等实践经验为依据而撰写的文献,也常被称为原始文献(不管以何种形式出现),具有原始性、创造性和分散性的特点。如期刊论文、专著、科技报告、学位论文、专利说明书等。一次文献所记载的知识、信息比较新颖、具体、详尽,具有创造性、实用性和学术性等明显特征,能反映科学研究的直接成果,同时也是数量最大、种类最多、使用最广、影响最大的一类文

献,是人们检索利用的重要信息源。

(3) 二次文献:二次文献是根据一次文献外表或内容特征,按照一定规则和方法进行加工、整理、浓缩、按一定顺序编排形成的有序化文献。由于一次文献发表分散、杂乱无序,难以掌握,给人们检索利用带来一定困难,特别是在“书富如海”的今天,一次文献的传递和利用障碍更大。因此,有关信息部门将一定范围、类型的一次文献尽可能全地收集起来,编制成各种二次文献,如各种目录、索引和文摘等,主要作用是提供一次文献线索,是检索和利用一次文献的重要工具。

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。

(4) 三次文献:三次文献是为了一定目的和需求,对大量一次文献信息中的有关内容进行综合、分析、提炼、重新组合而成的再生性文献。它通常是围绕某个专题,利用二次文献检索搜集大量相关原始文献,对其内容进行深度加工而成。三次文献源于一次文献但高于一次文献,不仅编排有序,而且赋予新的认识和见解,具有综合性高、针对性强、系统性好、知识更加成熟的特点,有比较高的使用价值。

三次文献可分为综述研究和参考工具书两大类型:

① 综述研究类:是在大量原始文献成果基础上,对科学技术的发展趋向进行分析研究,是综合评述的产物。属于这类文献的有综述、评论、评述、进展、动态等,这类文献是对已有成果加以评论、综述并预测其发展趋势的文献,可以使人们在短时间内了解其研究历史、发展动态、水平等,具有较高的实用价值。

② 参考工具类:是在大量原始文献内容反映的原理、定律、事实、方法、公式、数据及统计资料的基础上,筛选出稳定、可靠而有用的知识,编写成供查阅的工具书,如百科全书、大全、年鉴、手册、指南等。

综上所述,从零次文献到一次、二次、三次文献,是一个从不成熟到成熟,由分散到集中,由无序到有序,由博而约,由博而深的不同层次的文献加工过程。每一层次所含知识信息的质和量不同,所起的作用也不相同。零次文献是最原始的信息资源,是生成一次文献的素材;一次文献是文献基本形式,是最主要的信息资源,是人们检索利用的主要对象;二次文献是对一次文献的浓缩和有序化,是检索一次文献的工具;三次文献是将大量分散的一次文献归纳、按知识门类或专题重新组合,多从新的高度揭示相关文献内容,是人们查考事实、数据的工具,是掌握某专题新内容、新动态的重要信息资源。

3. 按文献出版形式划分

(1) 图书:是文献中最古老的一种出版类型,也是现代出版物中品种最多、数量最大的一种。图书是编著者在大量收集、整理知识信息的基础上,结合自己研究成果和工作经验,进行全面归纳、总结、深化的结果。它传播的知识比较成熟可靠,系统全面,是掌握一门学科、一个专题知识的基本资料。图书的装订形式比较讲究,完整定型,首尾衔接,结构严谨,自成体系。有封面、书名页、目次(目录)、正文、版权页、封底组成一个独立的整体,以册为单元,每册在 48 页以上(国际文献工作基础标准 ISO/DIS127/II)。图书出版周期较长,报道的知识信息有一定时差,因此,开展科学研究,了解科技发展趋势、方向、前景等,若想掌握某方面的最新知识信息,图书常常不

能满足需要,还须借助于其他类型文献。

图书分为一般性图书和工具书两大类,一般性图书包括教科书、专著、论文集、丛书等。工具书指词典、百科全书、年鉴、手册、指南、书目、索引、文摘等。

(2) 期刊: 又称杂志(Journals 或 Magazine), 是一种汇集了多位著者论文的连续出版物, 具有固定的名称, 统一的开本, 定期出版, 每期都有年、卷、期等序号。其特点是出版周期短、报道速度快, 内容广泛, 知识新颖, 信息量大, 能及时反映当前国内外科技水平和动向, 是人们了解科技信息、掌握学术动态和发展趋势, 启发创新思维的重要参考文献, 是重要的信息源(约占整个科技信息来源的65%~70%)。

核心期刊是指该学科所涉及期刊中, 刊载论文多、信息量大, 论文水平高, 能及时反映本学科最新研究成果、前沿状况以及发展趋势, 受该学科读者重视并有一定权威性的期刊。

报纸也应算作是一种期刊, 它是一种以刊载新闻和评论为主要内容的散页的定期出版物。它有固定的名称和开本, 有年、月、日期号和顺序号, 只是出版时间更短, 更快, 信息更及时, 内容更广泛。

(3) 特种文献: 是指图书、期刊以外, 出版形式比较特殊的一类文献。这类出版物介于图书与期刊之间, 似书非书, 似刊非刊, 品种多, 数量大, 内容广泛, 出版分散, 收集比较困难(有的不公开发表), 但现实性强, 参考价值大, 能从不同角度及时了解当前某领域的发明创造、动态、水平以及发展趋势, 是文献中不可忽视的重要部分。主要有如下几种:

① 科技报告, 也称研究报告或技术报告, 是关于某项研究成果的正式报告或某研究课题进展情况的实际记录。其内容专深具体, 叙述详尽, 数据可靠, 推理严谨, 报道速度快, 所涉及的一般都是尖端科学和最新科研课题, 能反映一个国家的科学技术成果、动向和发展水平, 有直接的借鉴作用。

② 专利文献, 是指在专利审批过程中所产生的官方文件以及相关的出版物, 包括专利说明书、专利公报、专利文摘、专利分类表等。专利说明书是专利文献的主题, 是该项发明的详细说明。专利文献内容新颖、详细可靠、实用、具体, 是了解某些技术领域内新发明、新设计的重要途径。

③ 会议文献, 是指国内或国际学术会议上宣读、交流的论文、会议报告、会议纪要等文献。这类文献学术性强, 往往代表着某一学科或专业领域的最新成就和最新研究课题, 反映当代国内或国际科学技术的发展水平和趋势, 是了解学术动态、研究进展的重要渠道。

④ 学位论文, 是高校或科研单位的研究生, 为获取学位而撰写的研究论文, 主要指硕士论文和博士论文。学位论文所探讨的问题比较深入, 多有创新, 比较系统详尽, 对于科学研究和撰写学术论文均有参考价值。

⑤ 政府及国际组织出版物, 是指各国政府部门及其所属的机构发表的文件和报告, 以及国际组织如联合国教科文组织、世界卫生组织等出版的文献。这一类的文献对于了解一个国家有关科技的方针政策、发展方向和国际标准、动态具有一定参考价值。

⑥ 其他特种文献, 还包括标准文献、技术档案、产品资料等。

1.2.3 现代文献特点

现代文献具有其自己的特点,主要有如下几点:

1. **数量大、增长快** 社会发展越快,知识产生越多、文献形成也就越多。特别是近代科技成果的大量涌现,使得文献数量急剧膨胀,每年以 18% 的速度递增。生物医学文献数量及增长速度更为明显,占各学科领域之首,生物医学文献占全部文献的近 1/6。大量的文献和快速的增长,给人们检索文献和利用文献信息带来一定难度。

2. **分布广泛、分散交叉** 现代科学的特点呈现高度的分化和高度的综合化,各学科之间相互渗透、相互交叉,综合学科、边缘学科、交叉学科不断涌现,学科之间的严格界限逐渐模糊消失,学科之间的相互联系逐渐加强,自然科学与社会科学之间相互汇流。同一篇论文会涉及多个专业,一种专业期刊会报道不同专业的文章,这种现象导致文献分布出现既相对集中又高度分散的现象,一个专业范围的文献只有 1/3 刊登在本专业刊物上,其余的专业文献却分散在大量非专业刊物上。这一现象提示人们在检索和利用文献信息时,除注意利用专业性的检索工具外,更要注重利用综合性检索工具,扩大检索范围,以免造成严重漏检现象。

3. **使用寿命缩短** 现代科技发展日新月异,每时每刻都有新的发明和创造,新理论、新方法、新技术、新成果不断取代原有的理论、方法、技术和成果,反映科学技术水平和成果的文献随之出现新陈代谢加快、老化加剧、使用寿命缩短的趋势。该现象表现为文献的半生期(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)的缩短,生物医学文献半生期为 3 年,物理学 4.6 年,化工 4.8 年,化学 8.1 年,冶金学 3.9 年,地理学 16 年……生物医学文献的半生期为 3 年,就是指正在利用的生物医学文献 50%,其出版年限不超过 3 年。文献半生期愈短,其老化速度愈快,使用寿命也就愈短。文献的老化速度与学科信息量的增长有关,信息量增长越快,文献的老化速度也就越快。

4. **出版类型多样化** 现代的出版物突破了单纯的纸张印刷形式,由于声、光、电、磁等技术的广泛应用,使现代文献的记录方式和载体形式发生了重大变化,形成了刻写、印刷、光电存储等多种生产方式。因此,除纸质印刷文献外,还出现了缩微型、声像型、电子型等多种载体的文献。非纸质型文献知识信息存储密度大、传递速度快,形象性强,功能上大大优于传统的印刷型文献,这些新型文献在文献中所占比例越来越大,发行量越来越多。

1.3 信息检索及其重要性

面对未来的世界,如果不掌握一定的信息检索和利用能力,将会陷入信息的困境。要了解信息检索,了解信息检索的意义和重要性。

1.3.1 信息检索定义及类型

从广义的角度讲,信息检索包括信息的存储和信息获取两个过程。信息的存储是指通过对大量无序信息的选择和收集、著录和标引等方法,根据信息的外表或内容特征,按一定方式组织编排,编制成各种各样信息检索工具和检索系统,使之成为有序化信息

集合的过程。信息的获取是根据特定需求,运用组织好的检索系统和检索工具,将特定信息找出来的过程。狭义的信息检索仅指信息的获取过程。检索,即查询、获取。

信息检索是利用一定的检索工具,以科学的方法,从众多的文献中找出特定需要信息的过程。

根据检索对象和检索内容的不同,信息检索可分为文献检索、事实检索和数据检索三种。

1. 文献检索 文献检索(Document retrieval),是查找特定相关内容的文献,如查找某一专题或查找某一著者的文献,再从查出的文献中获取相关的信息,这是最主要、最基本的检索。文献检索分为题录检索、文摘检索和全文检索。题录检索查到的仅是一系列与课题相关的文献线索,还需按题录提供的线索去找原文;文摘检索查到的除有文献的线索外,还有原文的摘要,阅读摘要可了解原文的主要内容,但若掌握全部内容还需查找原文;全文检索则是直接查到所需原文,通过阅读原文掌握其中信息。全文检索随着计算机存储介质容量的不断增大和网络文献的快速发展,愈来愈受到人们的欢迎。

2. 事实检索 事实检索(Fact retrieval),是以特定事实为检索对象,检索的结果是有关某一事物的具体答案,因此事实检索是一种确定性检索。如“胰岛素合成”的年代、意义是什么?什么是克隆技术?何人何时首次报道?主要研究成果是什么等。

3. 数据检索 数据检索(Data retrieval),是以特定的数据为检索对象,检索的结果是经过测试、评价过的各种数据,可直接用于比较分析或定量分析,是一种确定性的检索。例如查找各种统计数据、公式、图表、计算式、指数、常数、药物剂量等等属于数据检索范围。

上述三种检索没有本质的区别,只是检索要求不同。事实检索和数据检索是直接从文献中检出所需要的信息,而文献检索则是先检出包含所需信息的文献,进行阅读分析,进一步获取利用信息。

1.3.2 信息检索意义及作用

信息检索重要意义和作用有以下几方面:

1. 信息检索是打开知识宝库的钥匙 文献是存储知识信息的最基本形式,是学习和科学研究工作的物质基础,是信息的源泉,查阅文献是获得信息的重要渠道。但随着科学技术的飞速发展,新知识、新成果不断涌现,出现了所谓“知识爆炸”,作为知识载体的文献可以说“浩如烟海”,并且品种繁多,类型复杂,致使任何专家学者都不可能浏览所有的相关文献。面对文献海洋,查找其中自己特定需要的部分文献,如不懂文献检索,不掌握文献检索方法和技能,就会陷入“找不到,读不完”的困境,就会被文献海洋淹没。因此,信息检索成为科技人员必备的一项基本功。

“书山有路勤为径”,这是我国古人鼓励人们要勤奋学习的名言,但在知识经济时代,面对文献海洋,光靠“勤”已远远不够了,还要加上“巧”,这个“巧”就是掌握知识信息的技巧,信息检索就是最主要的技巧,能解决海量文献与特定需要之间的矛盾。学会文献检索,掌握了文献检索方法和技能,就能获取、利用文献中的信息,就等于掌握了打开知识信息宝库的钥匙,找到了一条吸收、利用大量新鲜知识的捷径,使自己能在知识的海洋里有目的地、准确地获取、吸收、利用所需知识信息。

2. 借鉴他人成果, 促进科研发展 科学研究具有继承和创造两重性, 任何科学技术的发展, 都离不开继承、借鉴前人经验和研究成果, 没有继承就不可能有发展, 没有继承就不可能有创新。牛顿说过: “如果我比别人站得高些的话, 那是因为我站在巨人肩膀上的缘故”, 这句话精辟地说明了继承在科学事业中的重要地位。科研活动必须在了解、掌握前人和他人已有成果基础上, 再通过自己创造性的思维和实验来进行, 也就是要在充分利用别人创造的知识的加速自己的发展。日本的发展就是一个典型的例子, 从二战后的战败国经过短短二三十年, 就能与欧美等发达国家并驾齐驱, 在某些方面甚至超过欧美, 成为世界经济强国、经济大国, 其主要原因之一就是他们善于吸收利用各国的信息资源, 在外国已有成果基础上再进一步发展, 这样起点高, 发展快, 既节省了大量投资, 又加快了发展速度。1967 年我国科学家首次成功合成胰岛素, 就是通过信息检索, 借鉴了英国牛津大学的 10 年研究成果, 在掌握了胰岛素蛋白质分子结构的基础上进行合成的, 这样就节省了 10 年重复研究时间, 处于该学科的领先地位。

科学研究是探求客观事物本质和客观规律的活动, 它的内容是观察新的现象, 发现新的规律, 创造新的理论、新的发明或新的产品。科研活动大体上分为三个阶段: 即论证、确定课题阶段, 实验、求证阶段和鉴定、发表阶段, 每个阶段都需要查阅大量文献获得有关信息。特别是论证、确定课题阶段, 这是关键性的第一步, 科研课题选定的恰当与否, 直接关系到其价值。首先要进行大量的信息检索, 通过查到的文献, 搞清楚自己所要选的课题是否有人研究过? 已有哪些成果? 国、内外水平怎样? 进展如何? 还有哪些空白以及发展趋势等等, 只有掌握了这些信息, 才能判断课题是否具有创新, 是否有研究价值, 再来决定取舍。如果科技人员信息意识淡薄, 不善于从文献中了解有关的科技动态和进展, 不注重捕捉和利用有用的信息, 就会造成研究项目或新产品的开发重复他人劳动, 走弯路, 造成人力、物力的浪费。上海某研究所科研人员利用 10 年时间, 研究成功的“以镁代银”做保温桶镀层的科研项目, 在鉴定时却发现英国在 40 年前就已研究成功, 并在多个国家申请了专利, 这就是情况不明, 盲目上马而造成的惨重教训。

我国信息素质教育起步较晚, 相当一部分科技人员信息意识不强, 检索信息的能力较差, 很多人知道信息重要性, 但不知道如何查找, 到何处查找, 致使一些科研项目出现低水平的或无效的重复。诺贝尔奖获得者杨振宁博士和一个日本代表团对我国科研项目做出过相同的评价: 有 40% 的项目与国外重复或已有成果。

信息检索是科学研究的重要组成部分, 科技人员在整个科研活动过程中必须查阅大量的文献, 参考文献越多, 思路越宽广, 获得成功的把握也就越大。但查阅文献所用时间较多, 据美国科学基金会和日本国际统计局初步统计, 查找文献时间占整个科研时间的 50.9% (开始思考时间占 7.7%, 实验研究占 32.1%, 编写科研报告占 9.3%)。如能掌握文献信息检索知识和方法, 及时有效地查到所需文献, 可以省时省力, 缩短整个科研周期。

3. 更新知识结构, 提供终生学习途径 随着社会的发展, 新知识不断涌现, 人们需要终生学习, 不断更新知识结构, 才能适应社会发展的需求。美国工程教育协会曾估计, 学校教育只能赋予人们所需知识的 20%~25%, 而绝大部分的知识是在走出校门后靠自己不断地学习去获得。高校是人才的摇篮, 培养出的学生既要有创造力, 又要有独立科研的能力, 既善于解决问题, 又富有创新精神, 这就要求大学生不仅要在学校掌握好教

师传授的知识，更重要的是要学会自己去获取知识。英国情报学家布鲁克斯(B.C.Brookes)提出关于信息作用于知识结构的方程： $K(S)+\Delta I=K[S+\Delta S]$ ，方程式中 $K(S)$ 为原有的知识结构，即一个人的背景知识； ΔI 为信息增长量，即能够理解、整合到自己知识结构中的信息； $K[S+\Delta S]$ 为吸收了新的信息后形成的信息知识结构。布鲁克斯的理论揭示：能够满足人们信息需求的不是所有的信息，而是使人们原有知识结构发生变化的那一小部分知识或信息 (ΔI)；这一小部分知识或信息的利用不仅仅是量的改变，而且是知识结构的改变，“改变”是信息需求满足的必然结果；一小部分知识或信息 (ΔI) 的信息需求满足状况在很大程度上取决于原有的知识结构 $K(S)$ 。

信息检索是打开信息宝库的钥匙，掌握信息检索方法和技能，是更新知识，形成合理知识结构最有效的手段，做到终生学习，不断进取，适应信息社会的发展需要。

(薛平安)

第2章 信息检索基础

信息检索系统包括信息的存储和查找两部分。信息存储是检索工具的编制,专职人员将大量分散的文献收集起来,根据一定的检索语言和方法进行标引、编排成有序的、系统化的检索工具。信息检索是检索系统和检索工具的利用,用户根据自己需要,利用一定的检索语言和相对应的检索途径,从检索工具中找到所需的文献信息。信息的存储是检索的基础,检索则是实现存储的目的和手段。

2.1 信息检索系统及检索语言

信息检索需要利用一定的检索系统和检索工具,不管检索系统和检索工具编制还是利用,都要遵循一定规则,使用统一的语言(检索语言),否则信息检索将无法完成。

2.1.1 信息检索系统

信息检索系统是指根据特定的信息需求而建立起来的一种有关信息收集、加工、存储和检索的服务工作系统,主要目的是为人们提供信息服务。

按照信息检索的实施手段,把信息检索系统分为手工检索系统和计算机检索系统。

手工检索系统又称传统检索系统,是用人工检索信息的系统,也是计算机检索系统产生的基础,主要利用各种印刷型目录、题录、文摘和参考工具书。利用手查、眼看、脑子想来实施和完成检索。特点是方便、灵活、判断直观、查准率高,可随时根据需求修改检索策略。但因该系统收录信息范围有限,更新速度慢,又全凭手工操作,检索速度慢,易造成漏检。

计算机检索系统是由手工检索逐步发展而来,是利用计算机进行信息的存储和检索的系统。计算机检索系统的特点是检索速度快、检索途径多、使用方便,检索词之间可灵活组配进行复合检索,因此,计算机检索已成为主要的信息检索方式。

2.1.2 信息检索原理

信息检索的原理,是通过对大量分散无序的文献信息进行搜集、加工、组织、存储,建立各种各样的检索系统,并通过一定的方法和手段使存储与检索这两个过程所采用的特征标识达到一致;以便有效地获得和利用信息源。

文献信息的存储是为了检索,是检索的前提和基础。存储的过程,主要对信息源进行标引,将其外表和内容的特征(如文献的标题、作者、来源和主题等)用特定的检索语言转化为一定的标识(如主题词、分类号和类目名称等),再将这些标识按一定的顺序编排后输入检索系统,从而提供有章可循的检索途径。为了保证文献信息能存得进、取