

XINXI
JIAN SUO
YU
LIYONG

信息检索与利用

和正荣 陈克巧 张 帆 雷桂萍 编著



重庆大学出版社



信息检索与利用

和正荣 陈克巧 张帆 雷桂萍 编著

重庆大学出版社

内 容 简 介

本书主要介绍:信息检索的基础知识、基本原理、基本理论;结合国内外检索工具典例,深入讲授检索语言与检索技巧;突出了国际联机、因特网及光盘等计算机检索的内容;新增了企业竞争信息的检索、加工整理、分析研究及研究成果写作等篇章,从而较全面地反映了这门课程的最新进展。

全书具有结构合理、重点突出、系统简明、取材新颖、理论联系实际、易于掌握和实用性强等特点,可作为理工类高等院校本、专科生和研究生的教材,也可作为高等院校、科研机构、企事业单位科技人员和图书、信息部门有关工作人员的参考用书。

图书在版编目(CIP)数据

信息检索与利用/和正荣等编著. —重庆:重庆大学出版社,2000.8

ISBN 7-5624-2226-5

I. 信... II. 和... III. 情报检索 IV. G252.7

中国版本图书馆 CIP 数据核字(2000)第 35150 号

信息检索与利用

和正荣 陈克巧 张 帆 雷桂萍 编著

责任编辑 曾令维

*

重庆大学出版社出版发行

新华书店经销

重庆电力印刷厂印刷

*

开本:787×1092 1/16 印张:14.75 字数:368 千

2000 年 8 月第 1 版 2000 年 8 月第 1 次印刷

印数:1—12 000

ISBN 7-5624-2226-5/G·309 定价:18.00 元

前 言

在几千年的演进与发展过程中,人类社会经历了原始社会、农业社会、工业化社会,随着20世纪末“信息高速公路”计划的实施,“信息2000年”的构想,人类正阔步进入信息社会。实际上人类自进入农业社会以来,无论人们有无意识,信息已经积极地发挥着重大的作用,他极大地推动了生产力的向前发展。特别是发展到20世纪后期,人类在饱尝工业经济硕果之后,开始步入知识经济时代,人们终于认识到,信息是与物质和能源可以相提并论用以维系人类社会存在和发展的三大要素之一。

必须强调的是,现在我们讲的信息已绝不是80年代以前我国图书、情报界所曾接受的那个概念及范畴,其内涵和外延已经发生了极大的变化。

首先,“信息”已不只是80年代以前人们熟知的以印刷文本为载体的信息,他的范围已大大扩展,增加了视像信息、机读信息、电子信息、多媒体信息、网络信息以及人类借助现代信息技术正在探索和识别的其他一些形式的信息。

其次,信息是消息,是生灵万物相互赖以生存和进化的讯息,是人类或任何接受体感知的客观存在和差异,是客观事物之间相互作用和联系的表征,是减少不确定性的一种客观存在和能动过程。

第三,信息是促进社会、经济和科学技术发展的一种新型资源。

第四,现代“信息”的概念与半导体技术、微电子技术、计算机技术、通讯技术、网络技术、多媒体技术、信息服务业、信息产业、信息经济、信息化社会、信息管理和信息论等含义紧密地联系在一起。

第五,过去与“信息”混用的“情报”一词,在我国《辞源》和《辞海》中指的是敌情的报告。50至80年代初,我国又把政府重要部门以及经济、金融、科技等要害部门的那些时间性强、严格限制于内部了解、使用和需要保密的情况、文件、报告等均视为“情报”。在社会主义建设时期“情报”一词又偏重于科技领域,现在我们知道的“情报”,就其内涵而言,除了少量确属机密性的真正“情报”外,大量的所谓“情报”正是国际上流通的“Information”(信息)。所以,在我国,“情报”包含于“信息”之中,“情报”是“信息”的下位类,同样,“文献学”、“图书馆学”也属于“信息”的一部分,都属于“信息”的下位类。他是中国社会、政治、经济和所处的环境的特征性决定的,是被信息学的理论与实践所证明了的不可改变的事实。

从以上众多的事实说明,仅用“文献检索与利用”或“科技情报检索与利用”等作为培养大学生信息工作能力的课程名称,已远远满足不了社会的需求。所以,我们将这门课程名称更名为“信息检索与利用”,而且从课程内容上作了多方面的调整。

《信息检索与利用》是根据当前信息化工作发展的趋势,在多年教学实践的基础上,吸取目前国内同类教材的优点,结合理工类高校实际编写而成。全书共分四篇:第一篇主要介绍信息检索与利用的共性知识、原理、方法及基本理论;第二篇结合理工类院校特点,精选部分国内外检索工具作为典例结合实际深入讲授检索理论,以便提高学生的检索技巧;第三篇旨在主要

采用现代化的检索手段,加强了国际联机、光盘及因特网(Internet)检索的内容,增加了上机检索课时及内容;第四篇介绍以竞争情报的检索为核心的信息加工整理、分析研究、研究成果写作、利用与信息服务的内容。所以本书既具有坚实的理论基础,又突出了先进的检索手段,同时增加了信息利用的内容,他是一本信息由手工检索为主转向以计算机检索为主的新型过渡性教材。

本书由昆明理工大学“文献检索与利用”课教研室和正荣、陈克巧、张帆和雷桂萍四位同志合作编写。其中第一、二、五章由张帆编写;第三、四、十、十一章由陈克巧编写;第六、七、八章由雷桂萍编写;第九、十二、十三章及附录由和正荣编写;何萍、罗浩莹、朱宁、张琨丽、徐勇五位同志参加了本书的部分工作。全书在共同讨论的基础上由和正荣、陈克巧负责统稿。

本书在编写过程中得到何萍等老师、云南省高校图工委文检课专业委员会、昆明理工大学教务处、图书馆、科研处和教材科的有关领导及工作人员的大力支持。书中还参考了图书、情报界同行和专家们的论著,并引用了部分论点,在此一并表示感谢。

限于编者学识水平,不足和错误之处难免。恳望同行和读者指正。

编者

2000年6月

目 录

第一篇 信息检索基础知识与基本理论

第一章 信息检索基础知识.....	1
1.1 信息源概述	1
1.2 信息检索	6
思考题	8
第二章 信息检索基本理论.....	9
2.1 信息检索载体	9
2.2 信息检索语言	12
2.3 信息检索原理及步骤	21
思考题	25

第二篇 手工检索工具典例

第三章 中文检索工具	26
3.1 中文检索工具概况及结构体系	26
3.2 中文检索工具的编排	27
3.3 中文检索工具的使用方法	33
思考题	35
第四章 标准文献及其检索	36
4.1 标准及标准文献	36
4.2 我国技术标准文献检索	38
4.3 国际标准及其检索	39
思考题	41
第五章 专利文献及其检索	42
5.1 专利基本知识	42
5.2 专利文献	48
5.3 国际专利分类法 (IPC)	50
5.4 中国专利文献及其检索	53
5.5 英国德温特 (Derwent) 出版公司的世界专利文献检索工具	56
思考题	68
第六章 美国《工程索引》	70

6.1 概况及总结构	70
6.2 EI 的主体结构	71
6.3 EI 主题词表及标题词语言	73
6.4 EI 的辅助结构	80
6.5 EI 的检索实例	82
思考题	84
第七章 英国《科学文摘》	85
7.1 概述	85
7.2 SA 文摘本文献条目的编排与著录	86
7.3 SA 索引本及索引体系	88
7.4 INSPEC 叙词表	91
7.5 SA 检索实例	93
思考题	96
第八章 美、英《金属文摘》	98
8.1 概述	98
8.2 MA 文摘本	98
8.3 MA 索引体系	101
8.4 叙词语言及冶金叙词表	102
8.5 MA 检索实例	106
8.6 合金索引	108
8.7 钢铁快讯	111
思考题	113
第九章 美国《化学文摘》	115
9.1 概况及特点	115
9.2 CA 文摘编排与著录	115
9.3 CA 索引系统及其使用方法	117
9.4 CA 检索实例	128
思考题	134

第三篇 计算机信息检索

第十章 计算机信息检索概论	135
10.1 计算机信息检索的原理及意义	135
10.2 计算机信息检索系统	136
10.3 计算机信息检索技术	138
10.4 联机信息检索	144
思考题	157
第十一章 INTERNET 网络化信息检索与光盘信息检索	158
11.1 INTERNET 概述	158

11.2	INTERNET 网上信息资源	164
11.3	INTERNET 信息检索工具与方法	167
11.4	INTERNET 信息检索策略与实例	169
11.5	光盘信息检索	171
	思考题	182

第四篇 信 息 利 用

第十二章	企业信息工作	183
12.1	信息的概念及分类	183
12.2	企业信息工作的地位、作用及特点	185
12.3	企业信息工作的内容	188
12.4	企业信息素材的加工整理与分析研究	195
12.5	企业信息工作的新模式	199
	思考题	201
第十三章	科技文献写作	202
13.1	科技文献写作的意义	202
13.2	科技文献的撰写	203
13.3	几种科技文献的写作	204
	思考题	214
附 录		215
参考文献		226

第一篇 信息检索基础知识与基本理论

第一章 信息检索基础知识

1.1 信息源概述

1.1.1 信息、知识、文献、情报

1. 信息

人类认识和理解“信息”的概念和内涵是近百年内的事情,但是作为一种客观存在,在远古的宇宙万物之中,在史前的猎、牧部族之中,在农业社会和工业化社会中,在当今的文明社会中,信息一直在积极地发挥着人类意识或没有意识到的重大作用。科学技术在 18、19 和 20 世纪所取得的空前进步,使人们终于认识到,信息是与物质和能源可以相提并论的用以维系人类社会存在和发展的三大要素之一。因此,只有科学地了解和认识信息,才能更好地把握信息,进而才能使信息更好地为科学技术、经济和社会发展服务。

那么,什么是信息(Information)呢?半个世纪以来,科学界一直在对信息的定义进行积极的探讨,这里列出一则具有广泛影响的定义:

“信息是指应用文字、数据或信号等形式通过一定的传递和处理,来表现各种相互联系的客观事物在运动变化中所具有特征性的内容的总称。”^①

信息被看成是物质的一种属性。可以认为,信息是对客观世界中各种事物的变化和特征的反映;是客观事物之间相互作用和联系的表现;是客观事物经过感知或认识后的再现。信息具有价值性、可传递性、可存贮性、可加工性、时效性等特征。

2. 知识

知识是人们通过实践对客观事物及其运动过程和规律的认识。人们在认识世界和改造世界的过程中,获得大量客观事物传递的信息,即感性认识或经验,然后对这些感性认识通过大脑进行加工处理,形成理性认识,即知识。知识形成的过程,就是人脑对客观事物传来的信息进行加工的过程,因此,知识不是大脑的自然产物,知识只能来自实践。

可见,信息包含了知识,知识是信息被认识了的部分。知识可分为主观知识和客观知识。在人类生活环境中,普遍存在着信息,信息经过人脑接收、选择、处理而形成知识,这种存在于

^① 刘昭东等. 信息工作理论与实践. 北京:科学技术文献出版社,1995

人类头脑中的知识称为主观知识。主观知识借助语言符号,通过各种物质载体记录下来,就变成可以传递的客观知识,即文献。

3. 文献

文献是记录有知识的一切载体。他是将人类知识用文字、符号、图形及声频视频信号等记录方式在甲骨、竹帛、纸张、感光材料、磁性材料等载体上记录下来而形成的。知识、载体和记录是构成文献的三个基本要素。知识决定文献的内容,载体决定文献的形态,记录则是构成文献的手段。

文献是传递知识的有效工具。文献是知识的一部分,知识的另一部分存在于人脑中或其他实物上。当文献中记录的知识,一旦传递到用户并为其所利用时,文献中的知识就转化为情报。当然,文献中的知识不可能全部转化为情报,转化的只是特定的一小部分,这部分知识是用户原来不知道的新知识,而且是能够解决实际问题的知识。

4. 情报

情报就是知识通过传递并起作用的部分,或者说是传递中的有用的知识。情报只是人类社会特有的现象,在人类社会存在之前无情报而言。把知识中的特定部分传递到需要这一特定部分知识的使用者那里,就产生了情报。知识性是情报的根本属性。但知识必须在特定的时间内针对特定的用户经过传递,并能为用户采用且有效益,这部分知识才可称为情报。传递性、效用性是情报的基本属性。可见,文献中记录的知识并不都是情报;而情报也不都是文献中记录的知识,因为情报还能以语言、信号等非文献形式存在。当然,知识转化为情报,还是主要通过文献这一传递方式。所以文献和情报二者关系十分密切,而且有所交叉。

综上所述,关于信息、知识、文献、情报之间的关系,可用图 1-1 表示。

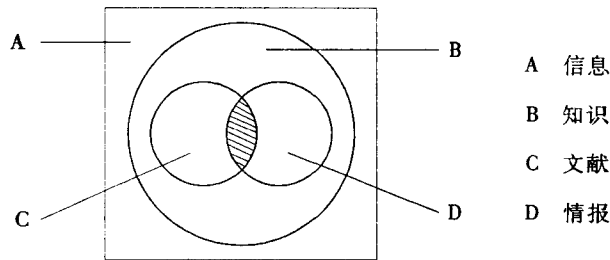


图 1-1 信息、知识、文献和情报关系图

1.1.2 信息源

人类对于自然界和社会的任何知识,无不来源于实践。科学研究人员、工程技术人员和一切参加物质资料生产的人们所从事的科技活动、生产活动及其成果,都是信息发生的源泉。

信息源按照其物质载体的不同可分为以下几种类型。

1. 人脑信息源

人脑信息源是指人们通过思维、交谈、讨论、发言、报告等形式不断沟通情况,发出并传递信息,也称口头信息源。其优点是:信息的传递直接、及时,针对性强,并且与信息接收者可以进行双向沟通,实现信息的及时反馈,对接收者来说信息的使用价值较大。但是,直接获取口头信息的机会是有限的,也不便为以后的研究进行信息积累。所以,在获取、传播、利用口头信息源的过程中,人们通常又把他转换成文献信息源。

2. 实物信息源

实物信息源是由产品展览、商品的陈列和展销等方式发出的信息来源,人们通过实地参观和考察和举办展览来加以交流传播。实物信息往往是直接为生产服务的技术情报,具有真实、直观、易检验、易仿制的特点,同引进技术和设计相比,花钱少且见效快,是值得重视的信息源。但是,实物信息只有经过复杂的分析和研究才能解析出来。

3. 文献信息源

文献信息源是一种经过加工的信息源,是将人类知识用文字、图形、声频、视频信号等手段记录在一定的物质载体上进行交流传播。文献信息源便于广泛传播、系统积累、长期保存和直接利用,是获取信息的最基本、最重要的来源。文献是以一定方式和手段将人类获得的知识记录在一定的物质载体上形成的,是最便于随时记录、阐明思想观点、广泛传播、系统积累、长期保存和直接利用的情报源。

科技文献是记录科技知识的一种载体,他累积着无数有用的事实、数据、理论方法、构思和假设,记载着无数的成功或失败的经验教训,反映着一定时代、一定社会条件下科学技术的发展水平,预示未来发展趋势和方向。科技文献是在空间和时间上积累和传播信息的最有效的手段,是获得科技信息的最基本、最主要的来源。

1.1.3 科技文献的作用

科技文献在人类社会实践中有以下的重要作用。

1. 存贮作用

科技文献是人类存贮并供人类分享利用的知识宝库,是人类社会科学技术得以发展的重要源泉。

2. 传递作用

科学技术的发展具有连续性和继承性。科技文献是科学技术得以继承、借鉴和发展的阶梯,是人类传递科技信息的主要手段。科技信息的传递有非正式传递渠道(如讨论、参观、演讲、展览等)和正式传递渠道(以文献形式传递),后者是最主要的传递手段。

3. 鉴别作用

科学文献是衡量学术水平和科研成就的重要标志。在评价一项科研成果时,往往是同前人已发表的有关文献进行对比,同已有文献中提出的观点、方法、工艺、数据等进行逐项比较,以证明该项成果的先进性或新颖性。

1.1.4 现代科技文献的特点

1. 数量庞大、增长迅速

随着科学技术的飞速发展,科技成果大量产生。作为科技成果表现形式之一的科技文献增长迅速,数量日益庞大,大约每10年就翻一番。

2. 载体形式多样

随着科学技术的飞速发展,仅仅通过传统的印刷型文献进行信息的交流和传递,无论从数量上、时间上或从交流服务方式上都已不能适应科学技术发展的要求。现代科学技术的发展,使记录知识的载体和手段日趋多样化。科技文献除印刷型外,还有缩微型、机读型、声像型、光盘型,而且,随着科学技术的发展,这些新形式载体的文献,在整个文献中所占比例不断增加。

3. 类型增加

除图书和期刊外,还有会议文献、科技报告、专利文献、标准文献、学位论文、产品资料等,而且每种文献都有特定的检索方法和各自的检索工具。

4. 语种扩大,译文增多

随着科学技术交流程度的不断增强,世界各国的科学技术得到普遍发展,科技文献的语种在扩大。为了克服语言障碍,便于科技文献的利用,大量翻译外国文献已成明显趋势,因而译文大量增多。

5. 内容交叉重复,分布分散

科学技术的迅速发展,使得传统的学科界限被打破,学科间相互交叉、渗透,使各专业文献内容交叉重复,布局分散。表现在:同一学科专业的文献分散于其他学科文献中;许多学科专业的文献分别涉及多门学科内容;同一篇科技文献可用不同出版类型和载体形式、不同语种重复出版。

6. 老化加速,失效加快

科学技术的飞速发展,使得新理论、新观点、新技术、新产品层出不穷,并迅速替代原有的科学技术,反映在科技文献上,就是科技文献老化加速,使用时间在缩短,迅速为新文献所代替。

1.1.5 科技文献的类型及使用价值

为把握利用科技文献的主动权,必须了解科技文献的类型及他们的使用价值。科技文献按不同的划分标准有不同的类型。

1. 按载体形式划分

(1)印刷型文献:他是以纸张为载体,以印刷(包括油印、铅印、胶印、复印等)为记录手段而产生的文献。这是一种传统的文献形式,其优点是便于阅读和传递;缺点是存贮密度低,体积大,分量重,保存期短,难以实现自动化信息检索。尽管如此,印刷型文献目前仍然占主导地位。

(2)缩微型文献:他是以感光材料为载体,以缩微照相为记录手段而产生的文献,包括缩微胶卷和缩微胶片。近年来采用激光全息技术,又产生了超级缩微胶卷。这类文献的特点是存贮密度高,体积小,节省空间,且不易变质,保存时间长,便于存贮珍贵文献。但这类文献不能直接阅读,必须借助复杂、笨重的缩微阅读机,尤其无法同时将多篇文献参照阅读。

(3)机读型文献:也称计算机阅读型文献,是以磁性材料为载体,通过编码和程序设计,用计算机可识别的机读语言记录和存贮知识的文献。这类文献存贮密度很高,存取速度快,可随时更新数据,促使信息实现自动、远程、快速传递和检索。但是该类文献必须借助计算机才能使用。

(4)声像型文献:是以感光材料或磁性材料为载体,以感光或电磁转换为记录手段而产生的文献,有唱片、录音带、幻灯片、录像带、电影胶卷等。他直接记录声音、图像,直观性强,容易接受,便于传播。

(5)光盘型文献:以高密度激光存贮技术为基础产生的文献。他将文献转换成激光信号记录在光盘上,阅读时,需通过光盘驱动器和电子计算机的配合。该类文献的最大特点是存贮密度非常大。光盘型文献可以看做是机读型文献,有时也可属声像型文献,但传统的机读

型、声像型文献相比,其记录手段不同。

2. 按出版类型划分

(1) 图书:图书是正式出版并具有一定篇幅的非连续性出版物。科技图书是对科研成果、生产技术和经验的概括和总结。图书的内容系统、全面,比较成熟和可靠,他是传播科技知识、教育和培养科技人才的主要手段,是综合、积累和传递科技知识的一种重要信息源。但由于图书出版周期较长,传递信息的速度较慢,因此其新颖性较差。当然,有些图书,作者在写作时曾参考和利用过一些当时的新资料也提出自己的新观点、新理论和新技术,但与其他文献类型相比信息含量仍然较少。

(2) 期刊论文:期刊又称“杂志”,是一种定期或不定期的连续出版物,长期使用固定刊名,每期均有连续的卷、期号和年月顺序号。期刊的出版周期短,刊载论文数量大,出版发行量大,内容新颖,影响面广。期刊论文一般都是作者研究的最新成果,通过期刊论文可以及时了解学科发展动态和趋势。据调查,科技工作者在科学研究工作中所利用的科技信息有 60% 来自期刊。由此可见,期刊论文是最重要的文献信息源,信息质量高,使用价值最大。

(3) 科技报告:科技报告是关于某项科学研究和技术开发最新成果或是阶段性进展情况的报告。科技报告分阶段性报告和终结报告。总的特点是每份报告单独成册,有机构名称,有统一编号,内容新颖专深,叙述详尽,其内容常经主管部门审定,故较成熟可靠。但科技报告中的阶段报告,审查不严格,出版发行又不规则,且具有一定保密性,因而流通受限制。科技报告是重要的信息来源,他在一定程度上反映一个研究机构乃至一个国家的科研能力和水平。美、英、法、德等国,每年都发行许多科技报告,其中最著名的是美国出版的四大报告,即美国军事国防部门的 AD 报告、美国商务出版局的 PB 报告、美国能源部的 DOE 报告和美国宇航部门的 NASA 报告。

(4) 会议文献:会议文献是指在学术会议上宣读或交流的论文。世界上每年举行上万次的学术会议来交流科研成果,讨论重大的学术问题。会议文献论题集中,针对性强,内容新颖专深,能反映某学科专业的最新进展情况。他是了解各国科技发展水平和动向的重要科技文献。会议文献常以图书、期刊等形式出版。

(5) 专利文献:专利文献内容较为广泛,这里讨论的专利文献主要是指专利说明书,即专利申请人为取得专利权向专利主管机关提供的该发明创造的详细说明书。专利文献内容详尽、新颖、实用,常常附有发明示意图。因为只有新产品、新技术、新工艺、新材料等的发明创造才有可能获得专利权,所以,专利文献对于工程技术人员来讲,是一种较为切合实际而又具有启发性的最新参考资料和信息源。

(6) 标准文献:是指对工农业产品和工程建设的质量、规格及其检验方法等方面所作的技术规定,是从事生产建设所必须遵循的技术法规。标准文献具有一定的法律约束力,标准文献有明确的适用范围及有效期,他是一个国家技术发展水平的具体反映。随着科技的进步,标准文献要不断修订和更新。

(7) 学位论文:是指为了获得各级学位而向学位授予单位递交的学术论文。学位论文分学士论文,硕士论文、博士论文三种,不同级别的学位论文质量和水平相差悬殊,一般来说学位论文论题单一,论述系统、详细,富有独创性,尤其是硕士论文和博士论文,对科技工作有一定的参考价值。学位论文一般不公开发表,由学位授予单位或指定的图书情报机构收藏。

(8) 产品资料:产品资料通常指产品样本、产品目录、产品说明书等,他是对定型产品的性

能、构造原理、规格、操作规程等的具体说明。产品资料形象直观、图文并茂,由于代表了已投产的产品,故其技术成熟、数据可靠,并附有较多的外观照片,因此他对新产品的造型、设计、技术改造、设备引进等都具有重要参考价值。产品资料一般由厂商免费赠送。

(9)报纸:报纸是有固定名称,以刊载各类最新消息为主的出版周期较短的定期连续出版物。报纸具有内容新颖、报道速度快、出版发行量大、影响面宽等特点,阅读报纸,是搜集最新科技信息的有效途径。但报纸受篇幅限制,报道内容不具体,不系统。

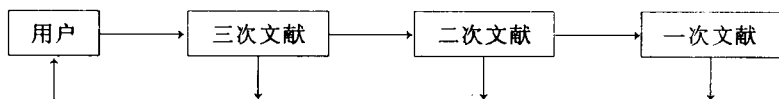
3. 按加工层次划分

(1)一次文献:作者以本人的研究成果为素材而创作的原始文献,不论撰写时是否参考或引用了他人的资料,也不论其载体和出版类型如何,都称为一次文献。一般地说,期刊论文、图书专著、科技报告、会议文献、专利文献、学位论文等,均属一次文献。一次文献论述比较详细而系统,能直接在科研、生产中起参考和借鉴作用,是科技文献的主体,是获得信息的主要来源。一次文献数量庞大,分布分散,不便于直接查找利用。

(2)二次文献:又称检索工具,是将大量、分散、无序的一次文献,用科学的方法进行加工、整理而产生的具有有序化、浓缩化特征的文献,如目录、题录、文摘、索引等。他从主题、分类、著者、题名等不同角度提供查找一次文献的线索。

(3)三次文献:指在利用二次文献的基础上,选用一次文献,经过综合研究和归纳分析形成的述评性或综述性文献。三次文献通常是围绕某个特定的课题,在利用二次文献基础上,搜集选用大量的一次文献,系统地综合和分析某学科、专业或专题的发展历史、已取得的科技成就以及发展趋势,内容十分概括,如综述、专题述评、手册、辞典和年鉴等。

从一次文献到二次文献再到三次文献,实际上就是一个由分散到集中,由无组织到系统化的加工、整理过程,也是一个把文献逐步进行浓缩化的过程。从文献信息检索的角度看,一次文献是检索的对象,二次文献是检索的手段和工具,三次文献是信息研究的成果。用户在检索应用时,他们之间的关系是:



1.2 信息检索

1.2.1 信息检索的概念和类型

信息检索(Information Retrieval)是指将信息(文献信息)按一定的方式组织、存贮起来,并针对用户的需要查找出所需信息的过程。因此从广义上讲,信息检索包含了信息的存贮和检索两个不可分的部分。而我们通常所讲的信息检索,是指狭义概念的信息检索,即从检索工具和检索系统中查找出所需信息的过程。

信息检索根据其检索对象的不同,可分为文献检索、数据检索、事实检索。其中,文献检索是三者中最基本、最主要的方式。

1. 文献检索(Document Retrieval)

查找出用户所需文献的线索或者原文的检索称为文献检索。例如,查找某一研究课题一定年限内的有关文献,或对项目发明创造进行文献查新,或从事新产品开发时需要查找有关最新研究动态等等,均属文献检索。文献检索是一种相关性检索,检索结果是文献线索(文摘、题录)一般要阅读文摘后才能决定取舍。文献检索主要是通过二次文献(检索工具)进行,如目录、题录、文摘、索引等。

2. 数据检索(Data Retrieval)

数据检索是指查找出用户所需特定数据的检索。例如查找某一数据、公式、图表、价格、某种物质的化学分子式、某种设备的型号与参数等,均属数据检索。它是一种确定性检索,某结果是所需数据要么有、要么无。数据检索主要利用各种字典词典、百科全书、年鉴、手册、名录等参考工具书进行,也可通过计算机网络查找大量的动态数据。

3. 事实检索(Fact Retrieval)

事实检索是指以特定的事实作为检索对象的一种检索。凡是对某一事物、事件、主题的事实情况进行查询均属事实检索,例如某个人或团体机构的基本情况,某一事件发生的时间、地点、过程等。他和数据检索一样,虽然也是一种确定性检索,但检索到的事实数据,必经过分析、对比研究的基础上,方能应用。

一般来说,文献检索是信息检索的基本检索,他要比数据检索和事实检索复杂而困难,主要通过检索工具达到检索目的;数据检索和事实检索是信息检索的派生检索,主要通过参考工具书来达到检索目的。

在实际工作中,往往把“信息检索”与“文献检索”混同使用,这主要是针对不同侧面而言。当强调检索的目的时,我们说“信息检索”,即通过各种检索系统查找出所需的信息;当强调检索的手段时,我们说“文献检索”,即从文献型检索系统中查找出所需的文献型信息。信息检索包含了文献检索,文献检索是信息检索中最重要的类型。

1.2.2 信息检索的意义和作用

1. 信息检索的意义

人类社会已逐步进入信息时代。前已述及,当今文献信息具有众多特点,以致有“文献污染”、“信息爆炸”之说。如今人们查找所需文献信息感到越来越困难,因此,要想从浩如烟海的文献信息中,及时、准确、全面地查找到所需的信息资料,不学习和掌握信息检索的知识和方法是不可能的。只有用科学的方法,有效的工具,才能打开文献信息宝库的大门。信息意识(人们对特定信息的识别、分析、判断、吸收的敏感性和自觉性)和通过信息检索获取信息的能力是影响科技工作者从事科学研究和技术开发能力强弱的重要因素,同时也是影响一个国家、一个地区、一个单位科研成果多少和学术水平高低的重要因素。正因如此,从1984年以来,信息检索知识的教育受到全国高等院校,特别是理工科院校的普遍重视,国家教育部曾数次发文要求加强文献检索课的教育。由此可见,文献检索的重要性受到普遍重视。

2. 信息检索的作用

(1)信息检索是打开人类知识宝库的钥匙。人类已有五千余年的文明史,历代流传下来的以及当今出版的众多文献是巨大的知识宝库,用科学的检索方法系统地开发利用这一丰富的文献信息资源,对发展科学技术具有重要作用。

(2)信息检索能使科技工作者及时把握科技发展的动态和趋势。随着科学技术的飞速发

展,新理论、新观点、新技术、新产品层出不穷。通过信息检索,科技工作者可以随时把握住科学技术发展的脉搏和全局,及时了解国内外最新科研成果。

(3)信息检索有助于拓宽知识面,改善知识结构。当今科学技术的迅速发展,知识老化现象日趋严重,个人获取的知识不可能是一劳永逸的。人的一生中,需要不断地进行知识更新,不断完善知识结构,以适应社会变革和科技发展的需要。

(4)信息检索能避免科研重复,能加快科研工作的进程。通过检索,及时搜集、查阅、分析、利用国内外文献,可以避免科研上的重复劳动,节省大量人、财、物力和时间。另据统计,在从事科学研究工作中,有近一半的时间是用于查阅文献信息资料,如果熟悉信息检索方法,就能大大节省查找文献资料的时间,从而加快科研工作的进程。

(5)信息检索能加强科技交流,促进技术合作。科技成果主要是通过文献的形式公诸于世,科技工作者的科研成果都是以文献的形式进行交流。因此,掌握信息检索技能,及时获取国内外文献信息日显重要。

思 考 题

1. 什么是科技文献? 科技文献有哪些主要作用?
2. 现代科技文献具有哪些特点?
3. 试述科技文献的类型及他的使用价值?
4. 试述科技文献检索的意义和作用。

第二章 信息检索基本理论

2.1 信息检索载体

2.1.1 信息检索载体的特征及职能

信息检索载体是指用以报道、存贮和查找文献信息的检索工具。

1. 信息检索工具的特征

(1)必须详细描述文献的外部特征和内容特征。所谓文献的外部特征是指文献篇名、著者姓名、文献出处等。所谓内容特征则是指文献的主题词、分类号、内容摘要等。

(2)每条描述记录(或称款目)都必须具有各种检索标识(描述文献外部特征和内容特征的专门用于信息检索的词、词组或代码)。例如主题词、分类号、著者姓名、文献序号等。

(3)全部描述记录科学地组织成一个有机的整体。

(4)能够提供多种检索途径。

2. 信息检索工具的职能

(1)报道职能。检索工具可以揭示某一时期某一范围的科技文献的发展状况。通过检索工具对科技文献信息的报道,可以了解到学科的历史、现有技术水平和未来的发展趋势。

(2)存贮职能。检索工具能把有关文献的内容特征和外部特征著录下来。成为可以识别的一条条文献记录。并按一定的次序排列组织起来,以便于查找各类科技文献信息。

(3)检索职能。检索工具能够提供一定的检索手段,使人们按照一定的检索方法,及时、准确、全面地查找出所需文献信息。

2.1.2 检索工具的类型

检索工具按照不同的划分标准,可以分为不同类型。

1. 按检索方法划分

检索工具可以划分为手工检索工具和机械检索工具两大类。手工检索工具指各种印刷型检索工具;机械检索工具主要指机读制品、缩微制品等。

2. 按收录文献的范围划分

检索工具可分为综合性检索工具、专业性检索工具、单一性检索工具三种。

综合性检索工具收录范围广,涉及多门学科。例如美国的《工程索引》。

专业性检索工具收录的范围仅限于某一学科领域。例如美国《化学文摘》、《金属文摘》等。综合性、专业性检索工具报道的文献类型常常是多样性的,如期刊论文、图书专著、会议文献、科技报告、专利文献等。

单一性检索工具只收录某一种类型的文献,学科范围可广可窄。如:期刊论文的检索工具有我国的《全国报刊索引》;专利文献检索工具有英国的《世界专利索引》等;另外还有查找会