

SCIENCE & TECHNOLOGY INFORMATION RETRIEVAL

科技信息检索

SCIENCE & TECHNOLOGY INFORMATION RETRIEVAL

饶安平 编著

四川出版集团 · 四川科学技术出版社

科技信息检索

SCIENCE & TECHNOLOGY INFORMATION RETRIEVAL

饶安平 编著

四川出版集团·四川科学技术出版社

图书在版编目(CIP)数据

科技信息检索/饶安平编著. —成都:四川科学技术出版社,2008.2

ISBN 978 - 7 - 5364 - 6247 - 2

I. 科... II. 饶... III. 科技情报—情报检索
IV. G252.7

中国版本图书馆 CIP 数据核字(2008)第 013076 号

科技信息检索

KEJI XINXI JIANSUO

编 著 饶安平
责任编辑 谢 伟 冯建平
封面设计 新飞扬工作室
责任出版 邓一羽
出版发行 四川出版集团·四川科学技术出版社
成都市三洞桥路 12 号 邮政编码 610031
成品尺寸 169mm × 239mm
印张 10 字数 200 千
印 刷 四川五洲彩印有限责任公司
版 次 2008 年 2 月成都第一版
印 次 2008 年 2 月成都第一次印刷
定 价 28.00 元
ISBN 978 - 7 - 5364 - 6247 - 2

■ 版权所有·翻印必究 ■

■ 本书如有缺页、破损、装订错误,请寄回印刷厂调换。

■ 如需购本书,请与本社邮购组联系。

地址/成都市三洞桥路 12 号 电话/(028) 87734081

邮政编码/610031

前 言

《科技信息检索》一书,是作者根据多年教学科研经验和信息咨询实践而编写的。近年来,电子资源不断涌现,互联网作为科技信息的载体,功能强大,使用方便,用户从网上检索科技信息较为顺畅,但采用纸本式检索工具检索文献却遇到很多问题。

世界著名情报科学家[美]兰开斯特在 20 世纪 70 年代末预言,2000 年无纸社会即将到来,然而纸本式文献因其具有历史久远,普及性广,通用性强的特点,至今仍在广泛应用。由此,以作者看来,前苏联世界著名情报科学家米哈依诺夫把科技文献视为科技信息交流基础的观点,符合当今科技信息检索的现实。

作者在实际教学工作中遇到的不管是高年级本科生、研究生,还是一些校外研究机构的科技人员,来咨询的多是国外纸本式检索工具中遇到的问题。本书作者视科技文献为科技信息交流的基础,在编写纸本式检索工具的基础上,也阐述了计算机信息检索。

本书的编写遵循精练、新颖、实用的原则,既考虑课堂教学又兼顾自学。特别是书中所取检索工具片段多是从各国检索工具上直接选取的,大大提高了本书的实用性。

科技信息检索尚在不断发展,作者编写水平也待提高,书中不当之处,如蒙指正,不胜荣幸。作者 E-mail:raoap@tom.com

饶海平

2008 年 1 月

目 录

1 信息检索概论..... 1	2.5.5 叙词语言 19
1.1 基本概念..... 1	2.5.6 《汉语主题词表》 20
1.1.1 信息利用的起源与发展 1	3 中文检索工具 21
1.1.2 信息的重要作用 1	3.1 题录式的检索工具 21
1.1.3 信息 2	3.2 文摘式的检索工具 23
1.1.4 知识 2	4 美国《工程索引》 30
1.1.5 情报 3	4.1 概况 30
1.1.6 载体 3	4.1.1 出版机构 30
1.1.7 文献 4	4.1.2 历年来的变化 30
1.1.8 五个概念的相互关系 4	4.1.3 报道文献概况 30
1.2 文献类型及其特点 4	4.1.4 <i>Ei</i> 近年来报道各国期刊 的比例 31
1.2.1 文献意义简论 4	4.2 <i>Ei</i> 的出版形式 31
1.2.2 文献类型 5	4.2.1 《工程索引年刊》 31
1.2.3 现代文献的一般特点 8	4.2.2 《工程索引月刊》 31
2 信息检索语言 9	4.2.3 《工程索引卡片》 31
2.1 信息检索语言的定义 9	4.2.4 工程索引检索磁带 ... 31
2.2 信息检索语言的特点 9	4.2.5 工程索引缩微胶卷 ... 31
2.3 信息检索语言的类型 9	4.2.6 工程索引光盘数据库 31
2.4 分类语言 10	4.3 <i>Ei</i> 报道文献的内容 31
2.5 主题语言 15	4.4 文摘的编排 32
2.5.1 定义 15	4.5 著录格式 32
2.5.2 标题词语言 15	4.5.1 期刊论文的著录格式 32
2.5.3 关键词语言(基本是自然 语言) 17	4.5.2 会议文献 34
2.5.4 单元词语言 19	4.5.3 技术报告 34
	4.6 索引及其使用 34

4.6.1 主题索引	34	5 美国《化学文摘》	47
4.6.2 年刊的主题索引	35	5.1 概况	47
4.6.3 月刊的主题索引	36	5.1.1 收录文件和报道文献量	47
4.6.4 著者索引	36	5.1.2 CA 的特点	47
4.6.5 著者所在单位索引	36	5.1.3 CA 历年来的出版情况	48
4.7 工程出版物索引	37	5.2 CA 文摘的编排	48
4.7.1 编码出版物索引——按缩写刊名排列	37	5.2.1 文摘内容	48
4.7.2 编码出版物——按编码排列	37	5.2.2 编排原则	48
4.7.3 新的和改变了编码的出版物	37	5.3 CA 文摘著录格式	49
4.7.4 编码出版物相互参照	38	5.3.1 期刊论文的著录格式	49
4.7.5 不包括会议文献的非编码出版物	38	5.3.2 新书(及声像资料)的著录格式	51
4.7.6 会议出版物	38	5.3.3 专利文献的著录格式	52
4.8 文摘号转换索引	38	5.4 各类文摘文献出处的识别标志	53
4.9 机构名称、字首缩写	38	5.5 参见项目著录格式	54
4.10 缩写字、单位和略语表	39	5.6 CA 索引系统	54
4.11 工程主题词表	39	5.6.1 索引系统	54
4.11.1 主标题词	39	5.6.2 各种索引及其使用方法	56
4.11.2 副标题词索引	40	5.6.3 检索实例	69
4.12 检索举例	41	5.7 检索实例	70
4.12.1 例1	41	5.7.1 例1	70
4.12.2 例2	43	5.7.1 例2	71
4.13 出版物名称和著者姓名的音(字)译	46	5.8 CA 主题索引的阅读规则	72
4.13.1 俄文	46	6 英国《科学文摘》	74
4.13.2 日文	46	6.1 概况	74
4.13.3 改写字母	46	6.1.1 出版情况	74
4.13.4 中国的港台、较老牌的刊名、人名——用威式拼音与汉语拼音对照表	46		

6.1.2 报道的内容	75	7.5 《速报》的检索方法	106
6.1.3 摘报原文的文献类型和文 种	75	8 专利文献	112
6.2 SA 的编排和文摘的著录格式	75	8.1 专利基础知识	112
6.2.1 SA 的编排方法	75	8.1.1 专利	112
6.2.2 文摘的著录格式	77	8.1.2 专利文献	113
6.3 SA 的索引	79	8.1.3 授予专利的条件	113
6.3.1 主题指南	80	8.1.4 检索专利文献的目的	114
6.3.2 主题索引	80	8.1.5 职务发明与非职务发明的 区别	114
6.3.3 著者索引	81	8.1.6 专利权人	115
6.4 小索引	82	8.2 专利文献的特点	115
6.4.1 文献目录索引	82	8.3 国际专利分类法简介	115
6.4.2 图书索引	83	8.3.1 概况	115
6.4.3 会议索引	83	8.3.2 IPC 的分类系统	116
6.4.4 团体著者索引	84	8.4 如何确定 IPC	117
6.5 附表	84	8.5 中国专利的检索	118
6.5.1 引用期刊表	84	8.5.1 例 1	118
6.5.2 引用期刊增补表	85	8.5.2 例 2	120
6.5.3 缩略语对照表	85	8.5.3 例 3	121
6.6 INSPEC 主题词表(叙词表)	86	8.6 如何在 Internet 上检索中外专 利	123
6.6.1 字母顺序表	86	8.6.1 中华人民共和国国家知识 产权局专利系统	123
6.6.2 词族表	87	8.6.2 美国专利与商标局专利全 文数据库	124
6.7 检索实例	87	8.6.3 欧洲专利数据库	126
6.7.1 例 1	87	9 计算机信息检索	129
6.7.2 例 2	90	9.1 概述	129
7 日本《科学技术文献速报》	95	9.2 计算机信息检索基础知识	129
7.1 概况	95	9.2.1 计算机信息检索应具备的 条件	129
7.2 《速报》各分册的沿革和内容	96		
7.3 《速报》的编排与文摘著录格 式	98		
7.4 《速报》的索引	102		

9.2.2 数据库的一般结构	130	9.5.2 《中文科技期刊数据库》——国内最大的综合性文献数据库	139
9.2.3 检索的一般方法	131	9.6 常用的外文数据库	143
9.2.4 光盘检索	132	9.6.1 EBSCO 全文数据库	143
9.2.5 联网检索	132	9.6.2 EBSCO 的检索方法	144
9.3 Internet	132	9.7 西文期刊目次数据库	148
9.3.1 IP 地址	132	9.7.1 系统简介	148
9.3.2 域名	132	9.7.2 CCC 的检索	149
9.4 Internet 上的信息、通过 Internet 获取的信息	132	附表一	150
9.4.1 Internet 上的信息 ...	133	附表二	151
9.4.2 通过 Internet 获取的信息	134	参考文献	152
9.5 常用的中文数据库	134		
9.5.1 CNKI 清华同方系列数据库	134		

1 信息检索概论

1.1 基本概念

信息,从字义上理解,信即信号,息即消息,通过信号带来的消息就是信息。信息是物质存在的反映。不同的物质具有不同的存在方式和运动规律,各自发出不同的信息。信息普遍存在于宇宙中。根据发生源的不同,信息一般可分为自然信息、生物信息、机器信息和人类信息四大类。在自然界,枝头绿芽报送春天到来是自然信息;细胞染色体的遗传基因是生物信息;电讯系统中的脉冲信号是机器信息;人类社会活动中的各种语言、文字、图形符号是人类信息。

1.1.1 信息利用的起源与发展

自古以来,人类通过自身感官从自然界接收万事万物发出的各种信息,藉以认识自然和改造自然,谋求自身的生存和发展。经过长期的社会实践,人类为了组织和管理社会生产和生活,调整人际关系,发展生产力,从被动地接受信息到自觉地搜索利用信息,又随着科学技术的进步,提高了信息的存储和处理能力。今天我们已经进入信息时代,这样的时代,现代化通讯、计算机网络和信息高速公路成了重要的标志。我们深感处处有信息,信息有妙用。

1.1.2 信息的重要作用

信息在人类社会和科技发展中虽然不是直接从事创造性发明,但它在促进社会进步、生产发展方面与从事创造性的科学研究工作一样,具有重要作用。它可以使人类创造的知识发挥更大的效用,犹如交通运输业那样,虽然不直接生产物资,但它对社会进步与生产发展的影响,都与物资产业部门创造产品一样重要,甚至它具有更重要的作用。信息的重要作用,突出地表现在三个方面:

一是为科研服务,促进科技发展。搞好信息工作,可以避免科研工作中的重复劳动,少走弯路。在科研工作中,选题十分重要。选题不仅要根据需要,而且要了解该课题是否曾有人做过,尽量利用别人已经取得的成果、经验、教训,这就要靠信息。

避免重复劳动就可以节省人力,节约时间和经费,这样就可以加快科研的进程。搞好信息工作,有利于科研人员从繁重的查阅文献资料中解放出来,缩短科研周期,早出成果。

二是为生产服务,加快经济建设的步伐。信息是一种重要的资源,是构成生产力最活跃的因素。我们善于开发利用信息资源,就能有效地促进经济和社会的发

展。一项适时对路的信息,可以带来一种新产品,或在贸易谈判中处于有利地位。技术经济信息有利于产品的更新换代,质量的提高,促进技术的进步和生产的发展;市场信息能提高全民经济生产的协调性。不难看出,信息可以在经济建设的各个方面发挥作用。

三是为决策提供科学可靠的依据。正确的决策,取决于对客观实际的了解,对未来行动及其后果的正确判断。而正确的判断又主要依赖于全面、及时和准确的信息。

随着社会的发展,信息对于管理决策的作用在不断增大,甚至有超过对科学研究作用的势头。人们通常把信息看成是制定计划、规划、方针政策、技术路线不可缺少的参谋;是了解国内外科技发展动向,掌握先进科学技术,赶超世界先进水平的尖兵;是加速科研进程,促进生产发展,繁荣经济,实现现代化的重要手段。

1.1.3 信息

世间万事万物无不向外界发出信息,以表明自己的存在和繁衍。不同的事物,有不同的特征。这些特征一定会通过一定的物质形式给人带来某种信息。我们通过识别自然界和社会的不同信息来区别各种事物,从而认识世界和改造世界。

因此,信息可定义为:事物运动状态的表征。事物以及具有自动控制系统的机器,通过感觉器官或相应设备与外界进行交换的一切内容,是表达其具体特征的信号。现在人们谈论的信息往往泛指现代信息,现代信息是指以电子技术为基础的信息。

哲学家认为信息是一种普遍属性,是物质存在的方式和运动的规律与特点。信息既不是物质,也不是能量,但是它与物质、能量有着密不可分的关系。首先,信息源于物质,源于物质的运动,没有物质也就没有信息。信息的产生、传送、存储、处理都以物质为依托。其次,信息也离不开能量,没有能量,信息收集、处理、转换、传递将无动力。同样,物质与能量都离不开信息,没有信息,物质和能量的存在形式、特征、运动状态等等就无法描述。

由此可见,物质、能源、信息确是当代社会的“三大支柱”。物质给人类提供材料,能源向人类提供动力,信息所奉献给人类的则是管理、控制物质和能源的知识和智慧。

1.1.4 知识

知识是人类对自然和社会运动形态与规律的认识和描述。知识是人类在认识世界、改造世界的过程中思想认识和实践经验的综合,是人类历史活动的结晶。人类通过信息感知世界、认识世界和改造世界,而且根据所获得的信息组成知识。

可见,知识是信息的一部分,是一种特定的人类信息。人类生活环境中普遍存在信息,信息是知识的原料,这些原料经过人脑接收、选择、处理,才能组合成新的系统化信息,即新的知识。这种新的知识存在于人脑中就是主观知识。一旦将头

脑中的认识结果通过某种物质记录下来,就变成了可以传递的客观知识。随着人类的深入发展,这种客观知识已逐步形成较完整的知识体系。人类的进步,就是知识的创造、积累和利用的过程。在科学技术是第一生产力的今天,知识已成为核心资源,知识也是经济增长的动力。

1.1.5 情报

情报是一种普遍存在的社会现象,在社会生活中,人们经常都在不同领域内自觉或不自觉地传递情报。然而如何科学地定义情报,则提法颇多。情报学学术界曾对情报的定义展开了非常热烈广泛的讨论。据1991年情报刊物的统计,国内外有关情报的定义多达129种。这表明情报仍是一个发展中的尚不够稳定的概念。在不同的历史时期,其社会功能和概念都有所不同。在古代,情报首先产生于军事领域,“战时关于敌情之报告,曰‘情报’”。这就是我国早期的情报定义,这种定义多带有谍报的含义。到上世纪90年代国家科委将科技情报改为科技信息,改名的目的之一就是以免把情报与谍报混淆。近代,随着现代科学技术的迅速发展,新知识大量涌现,印刷、复制技术,通信和交流技术,以及信息加工处理技术不断更新。控制论、信息论、系统论的相继创立;以电子计算机为中心的现代科学技术的广泛应用,更为情报的传递交流提供了崭新手段,进一步加深了人们对情报的科学认识。

综上所述,情报可定义为:情报就是作为人们传递着的准备发挥效用的知识。

一般认为情报具有三种基本属性:

一是知识性。情报的本质就是知识。没有一定的知识内容,就不能成为情报。知识是情报最主要的属性。

二是传递性。情报可以通过一定的物质载体和能量进行传递;要获得情报必须经过传递。通信技术的不断进步和日趋完善,情报服务的发展,网络信息目录服务的发展,都是为了改善情报的传递,这样才能满足社会对情报的需求。传递性是情报的基本属性。

三是效用性。人们生产、传递情报是为了向社会提供有使用价值的知识,人们选择、接收情报也是为了有效地利用已有的知识成果,两者在效用上的共同目的,就是实现了情报的价值和使用价值,这就是情报的效用性。情报与用户之间存在着密切的联系。情报为用户服务,用户需要情报,效用性正是用以衡量情报服务工作好坏的标志。

知识性、传递性、效用性是构成情报的三项基本属性。知识凝聚着情报的内核,传递推动着情报的运动,效用体现出情报的价值,三者共同描述了情报这一现象。

1.1.6 载体

载体是信息赖以存在的物质外壳。通常所说的载体指文献载体,即可供记录信息的一切人工附载物。古代文献载体有甲骨、简策、羊皮、纸张等。现代文献载

体的种类主要有纸张、磁性材料、感光材料等。随着科学技术的发展,文献载体的信息容量逐步扩大,保存日益完善,使信息交流的效能不断提高。

1.1.7 文献

文献是记录有知识的一切载体。这些知识采用文字、符号、图形、声频、视频等手段记录下来。文献是用于固定情报以及传递和使用情报的物质客体。文献是人们从事生产斗争和科学实验的记录,是人类精神财富中极其重要的一部分。它记录和反映着科学技术、社会发展的成就和水平,汇集着世世代代的广大劳动人民和科学工作者对事物认识的结晶,积累着无数有用的事实、数据、理论、方法,记载着无数成功或失败的经验教训,反映着人类的文明程度,是人类进步的重要基础。

1.1.8 五个概念的相互关系

信息、知识、情报、载体、文献这五者之间既有联系,又有不同。

信息的概念十分广泛。例如,宇宙射线给我们送来天体的信息。它是物质存在并运动着的表征。人们正是通过来自自然界和社会的不同信息以区别各种事物,并改造世界。

人类通过信息来认识世界,而在认识世界的过程中又把这种已经获得的信息,通过大脑思维重新组合,汇成知识。人类社会的进步,正是知识不断积累和更新的过程。

情报、知识都属于信息这个大的范畴,三者都可以被记录、存储、传递,而在一定条件下可相互转换。这三者是人类财富的根本源泉,有的学者把这三者称为:物、能、息(MEI: material energy Information)。

知识是系统化了的信息。一般的最新的知识通过交流、传递转变为情报,人们借助情报进行思维决策,又不断产生新的知识。

知识是一种基本描述的集合,它不管人们知道与否、理解与否;情报则是被理解的信息,是传递、交流着的知识。

知识可以是在遥远的年代形成的。知识可看作是信息通过不同的时间和不同的空间,由不同的人进行处理,最后形成的不断膨胀的综合体。

载体是信息的负载物,纸张是一种常见的载体,如果纸张不记载文字信息我们不一定把它看作载体,知识信息与物质载体具有不可分割性。知识信息一旦固化在载体上,我们就把它叫作文献。在信息时代,唯有充分开发和利用知识,才能启迪人类智慧,发挥巨大的力量,挖掘出各种财富的源泉。

1.2 文献类型及其特点

1.2.1 文献意义简论

文献是记载知识或信息的物质载体,全世界的科技成果都是通过科技文献来

传播的。文献情报是最基本最重要的情报源。作为记录人类知识的文献,每十年左右增长一倍,面对这么庞大的文献海洋,为了更好地检索利用文献,我们应对文献类型进行划分。

1.2.2 文献类型

文献类型可按载体划分,也可按生产传递层次划分,还可按出版形式划分。

1. 按载体划分的类型

(1) 印刷型文献。印刷型文献是以纸张作为存储介质,以印刷、复印等为记录手段的传统的文献形式。它用途最广,便于阅读、流传和保藏。印刷型文献包括油印、石印、铅印、胶印等的文献。

(2) 缩微型文献。这是以感光材料作为介质、以缩微照相为记录手段产生的文献形式,它包括微缩平片、缩微胶卷等,比印刷型文献的体积小、价格低,但查阅时需专门的阅读器。

(3) 机读型文献。这是一种磁性材料为存储介质,以计算机存贮为记录手段的新型电子文献。如各种数据磁带、磁盘、光盘等,其优点是存储密度高,存取速度快,但必须配置计算机。目前流行的光盘型文献,它是一种弱光记录和再现信息的海量存储型文献。

(4) 声像型文献。声像型文献以磁性材料或感光材料作为存储介质,以磁性记录或光学技术为记录手段直接记录声音、视频图像。它也叫视听资料或直感材料。这类文献有唱片、录音带、录像带、科技电影、幻灯片等。优点是直观、生动,但成本较高且不易检索。

(5) 实物型文献。实物型文献是一种非文字形式的文献。如各种文物、展品、样本、标本、仪器、设备,都负载着极其丰富的知识。

以上五种文献,印刷型文献仍为主流,它相伴人类几千年之久,在高科技的今天没有被淘汰,在 21 世纪也不会被淘汰。

2. 按产生与传递层次划分的类型

文献按对其产生与传递层次划分,有所谓一、二、三次文献的概念。

(1) 一次文献。通常由作者本人以自己的科研成果作为依据创作或撰写的文献均称一次文献。判定某一文献是否是一次文献,是根据它的内容性质,只要是作者根据自己的科研成果而发表的原始创作,无论是手稿、期刊论文、科技报告、专利说明书,也不论是铅印还是复制都是一次文献。这是对知识的第一次加工,主要特点是创造性。

(2) 二次文献。它是指文献工作者将大量分散的、无序的原始文献加以筛选,加工整理,按文献特征进行提炼、浓缩、简化、编辑成系统的工具性文献,如文摘、索引、书目等检索工具,专用于查找或评价原文。这是对知识的第二次加工,主要特点是有序化。

(3) 三次文献。凡是选用大量有关文献,经过综合、分析、研究而编写出来的文献,称为三次文献。它通常是围绕某个特定课题,采用科学方法,对文献内容进行深度的加工、浓缩而成。人们常把这类文献称为“情报研究”成果,如综述、专题述评、学科年度总结、数据手册等。三次文献的主要特点是综合性。

一次文献是检索对象(目标),二次文献是检索工具(手段),三次文献是情报研究成果,它既可作为检索目标,又可作为检索手段。

3. 按出版形式划分的类型

(1) 科技图书。科技图书的范围很广,从科普读物和各类教科书到深奥的专著等。它的特点是内容系统、全面、成熟、可靠。图书的缺点是出版周期长、报道速度慢。现在的图书都有一个 ISBN 号,即国际标准书号(International Standard Book Number),其定长为 10 位数字。ISBN 号是国际标准化组织 1972 年公布的一项通用的出版物统一编号方法。以《实用科技信息检索教程》为例:

ISBN7^①—5624^②—1099^③—2^④

①地域号(指国家、地区、语言区),例如:7 指中国、0 和 1 指英语国家,4 指日语;②是出版社(者)号;③是书(名)号;④是计算机校验码。

国际标准化组织(The International Standards Organization, ISO)规定,从 2007 年 1 月 1 日起,将国际标准书号(International Standard Book Number, ISBN)的长度从 10 位改变为 13 位。按照业界的协定,10 位 ISBN 写作“ISBN-10”,而 13 位 ISBN 写作“ISBN-13”。

(2) 科技期刊。科技期刊属于周期性出版物,具有品种多、数量大、报道速度快、内容新颖、能及时反映当前科技水平的特点。科技期刊是人们传递科技信息,交流学术思想使用的最基本、最广泛的手段。它是一类重要的文献信息源。外刊常以下面单词为名:期刊(Journal)、杂志(Magazine)、通报(Bulletin)、学报(Acta)、汇刊([拉]Transaction)。期刊也有国际标准连续性出版物编号:ISSN(International Series Standard Number)。ISSN 号全长 8 位,前 7 位是刊名代号,末位是计算机校验位。校验位数字 0~10,国内统一刊号 CN 为识别标志,以《情报杂志》为例 CN61-1167/G3:61—地区号;1167—序号;G3—分类号,定长 6 位。国际标准连续出版物编号以便实现对全世界期刊的科学管理。连续性和周期性,以及报道专题的多样化是期刊最显著的特征。期刊的文献特征是:①刊名和编辑单位不变;②定期出版,多为双月刊,其次还有月刊和季刊等;③有连续编号,如年、卷、期;④一般不再版和重印。

核心期刊:核心期刊是当前教育界、科技界谈论的热点话题,著名情报科学家严怡民教授早在 1983 年就明确定义为:“少数载有某一学科大量有关文献的期刊。”核心期刊表对图书情报机构选定期刊等很有参考价值;核心期刊本身不具备评价期刊论文水平的功能。

(3) 科技报告。科技报告是关于某项科研成果的正式报告,或是其进展阶段的实际记录。它的特点是每份报告单独成册、有连续编号、内容专深。科技报告是一次文献,许多最新研究课题和尖端学科,往往首先反应在科技报告中。科技报告的文献形式既像书又像刊,或既不像书又不像刊。

(4) 会议文献。会议录中收集的论文或报告均为会议文献。此外还有会前文献,它包括会议预印本、会议论文摘要等,但此类文献比会议文献更难获得。科技会议文献一般为一次文献,其特点是内容新颖、学术性强、能反映某专业国内外的最新发展水平。会议文献的形式主要为会议录(Proceeding),常以期刊专集的形式报道,也有以图书的形式出版,其常用的几种标志性题注有:Conference(会议)、Congress(大会)、Convention(大会)、Symposium(学术讨论会)、Workshop(专题学术讨论会)、Seminar(学术研讨会)、Colloquium(学术讨论会)、Meeting(会议)。

(5) 政府出版物。政府出版物是各国政府部门或机构出版的有关科学技术的政策性文件,它集中反映了各国政府有关部门的观点、方针、政策,对于了解某一个国家的科技和经济政策都有一定的参考价值。

(6) 专利文献。专利是用法律来保护技术发明产权的制度,它能巧妙地解决技术公开与保密这一对矛盾。专利的范围几乎囊括所有的技术领域。当专利申请案提出和批准时,即公布由发明人呈交的专利申请文件,阐明该发明的目的、用途、技术梗概和专利权限。专利文献包含了丰富的技术情报、经济情报、知识产权情报。因此,对知识创新和技术创新均有直接的借鉴作用。专利文献与其他类型的文献相比更具有法律色彩,其特点是格式化、条文化。

(7) 标准文献。标准文献是标准化工作的产物,是一种具有约束力的法律性文献,同时也是生产和科研的共同依据。标准文献对提高生产水平和产品质量,合理利用资源,节约原材料,促进科技发展都有着十分重要的作用。

标准文献按内容分为:基础标准、产品及零部件标准、原材料及毛坯标准、工艺及其装备标准、方法标准五种。按颁行范围分为:国际标准、国家标准、部颁标准、企业标准四个等级。按执行功能分为:强制性标准、一般性标准、推荐标准。

(8) 学位论文。学位论文是为了获得高等院校和高等教育机构中不同学位的人员提交的课题论文。学位论文根据学位制,有博士、硕士和学士论文之分。学位论文是经审查的原始研究成果,具有内容专一、阐述详细、比较系统的特点。学位论文一般不出版,只供复制。但一些检索刊物常以文摘的形式报道硕士、博士学位论文。若需原文可向学位授予机构索取。

(9) 产品资料。产品资料是各国厂商为推销产品而出版的一种商业性样本和产品说明书等,它能反映国内外同类产品的技术发展过程和发展动向等。产品资料的特点:技术成熟可靠、装潢美观、图文并茂。对技术人员的产品造型和外观设计有很大的参考价值。缺点是这类文献不易较全地收集。

(10)其他科技文献。报纸、技术档案、科技电影等。作为报纸上刊载的科技信息,主要是科技发展远景的展望。技术档案是某一工程项目的完整忠实的记录,它包括工程图纸、图表、图片、任务书、协议书、技术合同等,这些文献都具有一定的参考价值。

1.2.3 现代文献的一般特点

现代科学技术发展有两大趋势,一是微分化,即学科划分越来越细、内容越来越专业;新兴学科不断涌现。二是积分化,即学科之间互相交叉相互渗透,要取得任何创新成果,必须高度综合运用各科知识。由于科学技术的高度发展,科技文献作为科技知识的记录载体,也呈现出新的特点。

1. 数量庞大,增长迅速

据统计,目前全世界每年出版的科技图书约15万种,科技期刊约3.8万种,每年出版专利文献100万件,其他文献(会议文献、科技报告、技术标准等)几十万种。预计科技文献的数量每隔七八年增加一倍。

2. 形式复杂,文种多样

文献形式的复杂也是一个不可忽视的因素。同一篇文献可能在学术会议上宣读,然后编入论文集(会议录);有的又在期刊上发表,在期刊上发表之后,有的还收编入各种图书。另一种情况是同一作品出版有多种文种的版本,如众所周知的钱锺书的《围城》有中、英、法、日、俄各种版本。就科技文献而言,有的在日刊刊登的文献本身就译自德文,再把它译成中文又在中文期刊上发表;有的期刊同时刊载多个文种的内容等,这类情况是屡见不鲜的。

3. 内容广泛,分散交叉

随着科学技术在广度与深度上的不断发展,文献的内容十分广泛,涉及的学科颇多,从已发表的文献来看大到飞机、大炮乃至宇宙;小到别针、基因。各种文献彼此分散交叉的现象严重。科技文献有的只存贮手稿或只提供复印本。由于现代科学技术综合交错,彼此渗透,使得文献的专业性质也不十分固定;又由于同一文献往往由一种类型转为另一种类型,这样就造成了文献的分散交叉。

4. 代谢频繁,效用悬殊

科学技术的高度发展,使得一种理论、一项技术、一个产品,从发现、研究、试验到设计、生产之间的周期越来越短,因而导致科技文献代谢频繁。据调查,基础学科文献8~10年失效一半,技术文献寿命不足10年,多数在5年左右。

根据国内外多次文献利用调查统计发现,有约6%的文献被多次利用,有49%的文献仅被引用一次,而从未被引用过的无效文献多达35%。从文献内容方面分析,某一学科或专业引用指数高、论述主题直接的核心文献,相对集中于约5%的核心期刊中。

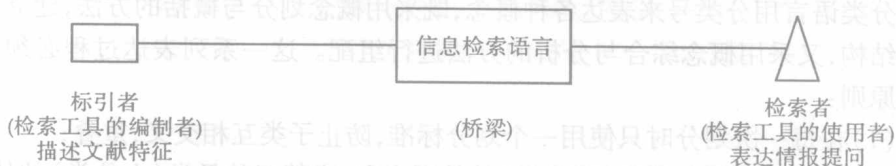
文献发展的上述趋势给文献信息的检索和利用提出了愈来愈高的要求。因此,如何在茫茫的文献海洋中去筛选获取所需的文献信息就成为信息检索的主要研究课题。

2 信息检索语言

2.1 信息检索语言的定义

信息检索语言,即信息检索领域中人们用来描述文献内容(主题)特征或外部特征以及情报提问特征的一种形式化语言,是根据信息检索的需要创造的人工语言,用于各种检索工具的编制和使用,并为检索系统提供一种统一的、作为基准的、用于信息交流的一种专用语言。因其使用的场合不同,检索语言也有不同的叫法。例如在存储文献的过程中用来标引文献,则叫标引语言;用来索引文献则叫索引语言;在检索文献过程中则为检索语言。检索语言把文献标引人员与检索人员联系起来,以便取得共同理解的语言。其用途是:作为描述文献特征,表达情报提问,使两者能互相沟通的依据。

标引人员:为文献编制标识的人。这里的标识就是揭示文献内、外特征的标签。



目前,世界上有成千上万种信息检索语言。例如:《汉语主题词表》《中国图书分类法》等。

信息检索语言应具备的基本条件:①具有语义和语法规则;②表达概念的唯一性;③易标引,易检索。

2.2 检索语言的特点

检索语言是由符号、代码或科技词汇组成。编制检索工具所遵循的“标准”,也是检索者遵循的“标准”。否则就很难使同一主题的情报表达与系统取得一致,就不能快速、准确地查找情报,或不能达到预期效果。检索语言在信息检索中起着重要作用,许多人把它叫做是标引者(文献加工人员)和检索者(用户)之间的桥梁。

2.3 检索语言的类型

世界上有成千上万种检索语言,按其原理可分为:分类语言、主题语言、代码语