



xinxijiansuo
yu
liyong

信息检索与利用

主 编 黄亚男

副主编 廖爱姣 谢岱 李鹏

 中南大学出版社
www.csupress.com.cn

信息检索与利用

主 编 黄亚男

副主编 廖爱姣 谢 岱 李 鹏

中南大学出版社

图书在版编目(CIP)数据

信息检索与利用/黄亚男主编. —长沙:中南大学出版社,2009
ISBN 978-7-81105-935-9

I. 信... II. 黄... 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(2009)第 109345 号

信息检索与利用

主编 黄亚男

☐责任编辑 谢贵良 陈海波

☐责任印制 汤庶平

☐出版发行 中南大学出版社

社址:长沙市麓山南路

邮编:410083

发行科电话:0731-88876770

传真:0731-88710482

☐印 装 长沙市神龙彩色印刷厂

☐开 本 787×1092 1/16 ☐印张 16.75 ☐字数 407 千字

☐版 次 2009 年 8 月第 1 版 ☐2009 年 8 月第 1 次印刷

☐书 号 ISBN 978-7-81105-935-9

☐定 价 28.00 元

图书出现印装问题,请与经销商调换

内 容 简 介

本书根据“信息检索”的新发展，系统阐述了信息检索的基本理论与方法、不同文献类型的检索工具及其使用方法；大型的综合性检索系统；国外几种在编排结构上具有代表性的著名综合性检索工具；计算机信息检索——联机、光盘、因特网和搜索引擎的相关理论与知识等。

本书既可以作为高等院校信息素质教育的教材，又是一本工具书，既可供普通读者查询信息时学习参考，也可作为信息专业人员推广信息资源的参考书。

前言

知识经济、信息时代的发展,网络技术的应用,使得信息量急速增长。面对浩瀚的、飞速更新的信息,如何有效地获取和利用有价值的最新信息,是现代科研人员必须充分意识和亟待解决的课题。

“知识的一半就是知道在哪里寻求它。”《信息检索与利用》课程的教学内容和理念也就是要培养大家在信息的海洋中找到知识的本领。进入网络时代后,数字资源和网络学术资源的检索成为检索课程教学内容的主要组成部分,其强大的检索功能使教学内容、教学手段进入了一个崭新的阶段。另一方面,由于信息素养教育在国内的不断发展,信息检索课程已成为高校开展信息素养教育的重要举措,和培养学生自主学习能力和独立研究能力以及创新能力的重要手段。信息素养的目标和内容要求人们具备:①正确确定所需信息的种类和范围;②高效检索信息;③正确地分析评价信息;④有效地利用信息进行创新,实现特定的目标,圆满完成学业或研究任务,不断提高自己的综合素质和解决实际问题的能力。

为了适应形势发展的需要,《信息检索与利用》将信息素养教育作为我们的基本指导思想,将最新信息检索方法与利用作为主要内容。我们根据近几年开设文献检索课程的教学实践,结合学院专业设置,以及各大检索系统的发展变化,编写了这本适用于各类本科生和研究生、科研人员的信息检索与教学,并能够反映最新信息检索进展的教材。在编写过程中,我们力求做到:①注重实用。《信息检索与利用》以各种常用的信息检索方法和技巧为主线,将信息检索的基础知识和理论与检索课实例相结合,着重介绍检索策略和技能,着眼于培养学生的信息素养能力。②内容新颖。尽可能吸收和介绍最新信息检索的新动态、新知识和新方法,使《信息检索与利用》能反映信息检索的最新进展,介绍一些特殊功能的使用,正确引导学生和科研人员掌握信息检索的真谛,提高信息素质水平,达到事半功倍的效果。③结构合理,体系完备。从信息检索的作用到信息检索原理、方法、步骤做了详尽的阐述,并对常用参考工具、国内外大型检索系统、大型检索工具、特种文献都做了介绍,其技术性、可操作性很强,便于读者掌握一定的理论基础和实践方法。

本书共12章,由黄亚男同志负责统稿,对初稿的个别章节、内容作了适当的调整和修改,并撰写第1章、第2章、第5章、第8章;廖爱姣同志撰写第3章、第6章、第7章;谢岱同志撰写第4章、第11章、第12章;李鹏同志撰写第9章、第10章。谢岱同志对所有英文进行了核对。图书馆汪文勇同志为我们检查核对了所有的网站及链接。

本书的编撰和出版得到了湖南文理学院图书馆、芙蓉学院的大力支持;得到了中南大学图书馆馆长张曾荣教授,中南大学周春山教授、张胜华教授的大力帮助,在此一并表示诚挚的谢意!

限于编者的学术水平,书中可能有所疏漏,欢迎业界同人指正!

编者

2009.5.28

目 录

第1章 绪 论	(1)
1.1 引论	(1)
1.2 信息资源及其类型	(4)
第2章 信息检索基础	(9)
2.1 信息检索原理	(9)
2.2 信息检索系统	(11)
2.3 信息检索方法	(14)
第3章 常用工具书	(19)
3.1 工具书概述	(19)
3.2 检索工具书	(20)
3.3 参考工具书	(27)
第4章 国内外大型综合性检索系统	(43)
4.1 CNKI 数字图书馆	(43)
4.2 超星读秀知识库	(46)
4.3 万方数据资源系统	(50)
4.4 中国高等教育文献保障系统(CALIS)	(56)
4.5 国际联机检索系统(DIALOG)	(61)
4.6 联机计算机图书馆中心(OCLC)	(65)
第5章 事实与数据检索	(70)
5.1 概述	(70)
5.2 名词、术语的检索	(70)
5.3 人物、机构的检索	(73)
5.4 统计资料的检索	(79)
5.5 数据、图谱的检索	(83)
第6章 图书信息检索	(86)
6.1 图书概述	(86)
6.2 联机公共目录查询系统(OPAC)	(89)
6.3 电子图书数据库	(105)

第7章 期刊信息检索	(118)
7.1 期刊概述	(118)
7.2 国内电子期刊数据库	(122)
7.3 国外电子期刊数据库	(140)
第8章 特种文献的检索	(154)
8.1 学位论文的检索	(154)
8.2 会议文献的检索	(158)
8.3 科技报告的检索	(162)
8.4 专利文献的检索	(167)
8.5 标准文献的检索	(172)
第9章 国外重要检索工具	(178)
9.1 美国《化学文摘》CA	(178)
9.2 美国《工程索引》EI	(185)
9.3 美国《科学引文索引》SCI	(190)
9.4 英国《科学文摘》SA	(195)
9.5 美国《生物学文摘》BA	(201)
第10章 搜索引擎	(206)
10.1 搜索引擎概述	(206)
10.2 常见中文搜索引擎	(210)
10.3 常见外文搜索引擎	(218)
第11章 网络免费学术资源的检索	(224)
11.1 网络免费学术资源简介	(224)
11.2 各类型网络免费学术资源的获取与利用	(226)
11.3 互联网上免费学术资源的开放获取	(234)
第12章 毕业论文的写作指导	(238)
12.1 毕业论文概述	(238)
12.2 毕业论文的选题	(240)
12.3 毕业论文的写作	(244)
12.4 毕业论文的基本要求和格式规范	(247)
参考文献	(252)
附录:《中国图书馆图书分类法》简表	(255)

第1章 绪论

在人们的日常工作、生活、学习、科学研究中无不包含着丰富的信息,快速准确、及时有效的检索和利用信息,是知识经济和网络时代对信息检索提出的新要求,也是走进新时代的人们必须具备的基本信息素养。我们很有必要熟悉掌握信息的一些基本概念以及信息资源、信息检索的内容与方法。

1.1 引论

1.1.1 基本概念

(1) 信息(information)

各种文献中有许多对于信息的不同理解和表述,其中最值得注意的是以下几种。控制论的创始人维纳(Norbert Wiener)认为:信息就是信息,既不是物质也不是能量。这个论述第一次把信息与物质和能量相提并论。信息论的奠基者——美国科学家克劳德·申农(Claude E. Shannon)认为:信息就是能够用来消除不确定性的东西。这是他1948年在著名论文《通信的数学理论》中提出来的,并成为信息论诞生的标志。这个论述第一次阐明了信息的功能和用途。比较流行的另一种说法认为:信息是事先不知道的报道。还有,哲学界认为:信息是事物普遍联系的方式。

早在一千多年前,南唐诗人李中在《碧云集·暮春怀故人》一诗中就留下了“梦断美人沉信息,目穿长路倚楼台”的佳句。当时,“信息”指的是音信、消息。

《科学技术信息系统标准与使用指南——术语标准》中对信息的表述:“信息是物质存在的一种方式、形态或运动形态,也是事物的一种普遍属性,一般指数据、消息中所包含的意义,可以使消息所描述事件中的不确定性减少。”

作为科学术语,从广义上讲,信息是事物运动时发出的信号所带来的消息,是事物存在方式和运动规律的一种表现形式。狭义的信息是指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。

总的来说,信息是用文字、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客观事物在运动变化中所具有特征内容的总称。因而我们可以这样认为,信息是事物存在的方式、形态和运动规律的表征,是事物具有的一种普遍属性,它与事物同在,存在于整个自然界、人类社会和思维方式中。

因此,信息本身是可以传递、扩散、复制、共享、增值的。信息、材料、能源一起成为人类生活的三个基本要素。材料、能源提供的是具体的物质,而信息提供的是知识和智慧。由此形成的信息科学就被人们确认为现代技术的三大支柱之一,它的发展大大地推动了人类文明的进步。

信息具有如下特征:

① 可识别性。信息是可以识别的,识别又可分为直接识别和间接识别,直接识别是指通过感官的识别,间接识别是指通过各种测试手段的识别。不同的信息源有不同的识别方法。

② 可存储性。信息是可以通过各种方法存储的。

③ 可扩充性。信息随着时间的变化,将不断扩充。

④ 可压缩性。人们对信息进行加工、整理、概括、归纳就可使之精练,从而浓缩。

⑤ 可传递性。信息的可传递性是信息的本质特征。

⑥ 可转换性。信息可以由一种形态转换成另一种形态。

⑦ 特定范围有效性。信息在特定的范围内是有效的,否则是无效的。

(2) 知识(knowledge)

与信息密切相关的一个概念是知识,即人类对于客观世界的认识。我国对知识的定义一般是从哲学角度作出的,如《博弈圣经》中是把“识别万物实体与性质的是与不是”定义为知识。《中国大百科全书·教育》中“知识”条目是这样表述的:“所谓知识,就它反映的内容而言,是客观事物的属性与联系的反映,是客观世界在人脑中的主观映象。就它的反映活动形式而言,有时表现为主体对事物的感性知觉或表象,属于感性知识,有时表现为关于事物的概念或规律,属于理性知识。”从这一定义中我们可以看出,知识是主客体相互统一的产物,它来源于外部世界,所以知识是客观的。但是知识本身并不是客观现实,而是事物的特征与联系在人脑中的反映,是客观事物的一种主观表征。知识是在主客体相互作用的基础上,通过人脑的反映而产生的。

人们为了进行知识的传递和交流,必须使知识能表现为人的感觉器官所感知的形式,即借助于文字、语言、符号、代码、电磁波、图像和事物等,这种表现形式就是信息。知识是人类通过信息对自然界、人类社会以及思维方式与运动规律的认识和掌握,是人的大脑通过思维重新组合的系统化的信息的集合。因此,人类要通过信息感知世界、认识世界和改造世界,而且还要根据所获得的信息组成新的知识。由此可见,知识是人类在改造客观世界中所获得的认识和经验的总和,是信息的一部分,是一种特定的人类信息。

世界经合组织(OECD)在1996年出版的《以知识为基础的经济》的年度报告中将知识分为四大类:

- ① know-what, 关于叙述事实方面的知识;
- ② know-why, 关于自然原理和规律方面的知识;
- ③ know-how, 关于技能和能力方面的知识;
- ④ know-who, 关于产权归属的知识。

获取上述知识的途径有两种:一是直接来源于产生信息的客观事物;二是通过信息载体或媒介。所获信息能否转化为知识,则取决于接受主体的认知能力。

(3) 情报(intelligence)

情报是外来语,来自日语,主要指“信息、资讯、消息”。情报是指被传递的知识或事实,是知识的激活,是运用一定的媒体(载体),越过空间和时间传递给特定用户,解决科研、生产中的具体问题所需要的特定知识和信息。情报具有3个基本属性:一是知识或信息;二是要经过传递;三是要经过用户使用产生信息。情报不仅取决于情报源,也取决于情报用户。也就是说,情报是在传递中对解决某一问题具有实际效用的知识或信息。情报是为未来的工

作服务的计划、决策、措施,是可以传递的,而且必须是具有针对性的知识,不一定是新的知识或信息。

(4) 文献(document/literature)

在我国古代,“文献”一词主要是指文字资料和言论资料以及阅历丰富、满腹经纶的贤人。随着历史的发展,文献的概念逐渐衍化为专指有价值的各个学术领域的各种文档资料,而原来含有的“贤人”一义则逐渐消失了。由于现代科学技术的发展,出现了各种各样的载体材料,发明了各种各样记录知识信息的方式,使得文献概念的外延不断扩大。载体从甲骨、竹简、羊皮、石板、丝帛、纸张发展到利用感光介质、磁性介质,文字也不再是表达思想的唯一手段,图形、音频、视频等技术手段同样成为表达思想、传递信息的重要手段。

人们通常把文献理解为图书、期刊等各种出版物的总和。国际标准化组织《文献情报术语国际标准》(ISO/DIS5217)中描述:“文献是在存储、检索、利用或传递记录信息的过程中,可作为一个单元处理的,在载体内、载体上或依附载体而存储有信息数据的载体。”《中华人民共和国国家标准·文献著录总则》(GB3792.1-83)定义:“文献是记录有知识的一切载体。”

由此可见,文献是由知识、载体和记录三个基本要素所构成。我们可以认为,文献是记录人类知识的一种载体,或理解为固化在一定物质载体上的知识。也可以理解为古今一切社会史料的总称。

总之,文献是记录、积累、传播和继承知识的最有效手段,是人类社会活动中获取信息的最基本、最主要的来源,也是交流传播信息的最基本手段。正因为如此,人们把文献称为信息工作的物质基础。

1.1.2 信息与知识、情报、文献的相互关系

(1) 信息与知识

信息是反映事物运动的状态及其变化方式,知识则是研究事物运动的状态及其变化方式的规律。只有通过信息的加工才能获得知识。如果没有信息,也就根本谈不上知识。知识是信息升华的结果,是浓缩的系统化了的信息。学者们把信息与知识的关系比做两个大小不同的同心圆,信息是外层圆,知识是内心圆,即表明了信息是知识的基础,而知识是信息的核心。

(2) 信息与情报

信息是事物存在的形式和运动状态,即存在于人类社会、自然界和人的思维活动中,其概念的内涵要比情报宽泛。情报是信息的一部分。信息与情报是包含与被包含的关系,情报必是信息而信息不一定是情报。

(3) 知识与文献

现在使用的“文献”一词,已经变成了既可囊括图书报刊,又可涵盖一切声像资料、电子出版物的超级词汇。凡是具有内容知识性、载体实在性,有记录方式、能够传播、便于积累继承基本特征的事物,都可成为文献。然而,知识与文献之间不能画等号。我国学者吴慰慈认为“信息的范畴远大于知识,知识是进入人们视野中的信息,是已被人们所感知与确认的信息,也有许多原始信息尚未被划入知识的领域”。因此,知识应该是信息的子集,而文献信息是以符号系统记录在文献载体上的信息。还有学者认为“知识是系统化的信息,而记载于

文献上的信息并非是系统化的信息”。所以，文献应该属于客观信息，是信息的子集。

概括起来，世界是物质的，物质运动便产生了信息；各种信息经过人们系统化的加工转化成知识；知识经过系统化的加工处理转化为情报；情报应用于实践，解决实践中存在的问题，创造出财富，即转化为生产力，进而产生新的信息。只有将自然现象和社会现象的信息上升为对自然和社会发展规律的认识，这种再生信息才构成知识。而情报是传递着的有特定效用的知识。知识信息被记录在载体上，形成文献。文献与知识是不同的

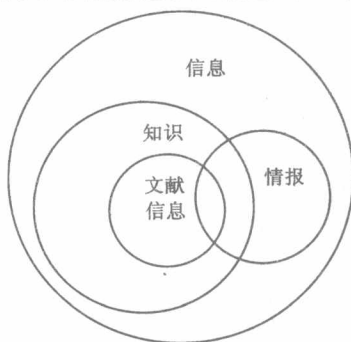


图 1-1-1 信息、知识、情报、文献信息的关系

概念，又有密切的联系。文献必须包含知识内容，而知识内容只有记录在物质载体上，才能构成文献。文献经过传递、传播、应用于理论和实践而产生信息。同时表明，信息包含知识、知识包含情报，它们之间不仅存在包含关系，而且可以相互转化。

1.2 信息资源及其类型

1.2.1 信息资源的含义

(1) 信息资源的含义

现在，人们已经普遍认为，信息资源是影响社会发展和进步的一种重要战略资源，对于信息资源的含义有广义和狭义之分。

柯平等在《信息管理概论》中指出狭义的信息指信息本身或信息的集合，准确地说是指信息内容；而广义的信息资源是信息及其相关因素的集合，除信息本身外，还包括与其紧密相关的信息设备、信息人员、信息系统、信息网络等，涉及信息的生产、分配、交换（流通）、消费等过程。

但很多人具体使用的是狭义的信息资源概念，即信息资源是指人类社会经济活动中经过加工处理有序化并大量积累起来的有用信息的集合。

信息资源是生产及管理过程中所涉及的一切文件、资料、图表和数据等信息的总称。它涉及到生产和经营活动过程中所产生、获取、处理、存储、传输和使用的一切信息资源，贯穿于生产管理的全过程。但它又不同于其他资源（如材料、能源资源），是可再生的、无限的、可共享的，是人类活动的最高级财富。

信息同能源、材料并列为当今世界三大资源。信息资源广泛存在于经济、社会各个领域和部门，是各种事物形态、内在规律和其他事物联系等各种条件、关系的反映。随着社会的不断发展，信息资源对国家和民族的发展，对人们工作、生活至关重要。因此，对于信息资源的开发和利用是整个信息化体系的核心内容。

(2) 信息资源的特点

信息资源与自然资源、物质资源相比，具有以下几个特点：

①能够重复使用，其价值在使用中得到体现；

- ②信息资源的利用具有很强的目标导向,不同的信息在不同的用户中体现不同的价值;
- ③具有整合性。人们对其检索和利用,不受时间、空间、语言、地域和行业的制约;
- ④它是社会财富,任何人无权全部或永久买下信息的使用权,它是商品,可以被销售、贸易和交换;
- ⑤具有流动性。

1.2.2 信息资源的类型

信息是一种重要的资源,这已成为人们的共识。对信息资源进行组织、管理、开发、利用也已成为人们普遍关心的问题。信息资源是在国民经济和社会信息化过程中,有利用价值的、数字化、网络化的信息内容。狭义地讲,信息资源是按照特定目的和规则,通过人们一系列采集、加工等劳动和创造过程之后以数字化形式储存在特定载体上可供利用的信息集合。信息资源涵盖了传统沿用的文献、情报、知识、数据等概念。从信息表达的内容看,信息资源既包括了经济信息、科技信息、教育信息、军事信息、文化信息等社会人文信息资源,也包括空间地理信息、气象信息、动植物物种信息以及人类遗传信息等自然信息资源。从信息内容的表达方式看,信息资源具有文字、数值、声音、图像等多种存在形式。

1.2.2.1 按信息源的表现形式分类

(1) 文献型信息源

文字是人们实现信息交流、通信联系而创造的一种形象符号,文献型信息源是以文字形式存储于各种不同载体上的信息,是目前内容最丰富、使用频率最高的信息源。

(2) 数据型信息源

是指以数值数据形式存储于各种不同载体上的信息,如统计数据、测量数据、理化数据,等等。

(3) 声像型信息源

是指以声音或图像形式出现的信息源,它比文字直观,易于理解,如电影、电视、CA 光盘、VCD 光盘等。

(4) 多媒体信息源

“多媒体”一词译自英文“multimedia”,媒体(medium)原有两重含义,一是指存储信息的实体,如磁盘、光盘、磁带、半导体存储器等,中文常译作媒介;二是指传递信息的载体,如数字、文字、声音、图形等,多媒体信息源是集声音、文字、图像、数据等多种通信媒介为一体的信息,一般以网络形式或光盘出现。

1.2.2.2 按信息的载体形式划分

(1) 印刷型文献

是指通过手写、印刷、复印等手段把信息固化在纸张上而产生的一种传统信息载体形式,如图书、期刊、报纸、档案、专利说明书、产品说明书等。

(2) 缩微型文献

是指以印刷体为母本,采用感光材料为存储介质,通过缩微技术而产生的一种信息载体形式,如缩微胶卷、缩微平片、缩微卡片等。

(3) 声像型文献

它是一种非文字形式的文献,以电磁材料通过特殊设备将声音、图像、动画等记录下来,

给人以直观、形象的感受。它包括唱片、录音带、录像带、电影电视片等。

(4) 电子型文献

采用高技术手段,把信息存储在磁盘、磁带或光盘等载体中,通过计算机对电子格式的信息进行存取和处理。电子出版物内容丰富,类型多,包括电子图书、电子期刊、电子报纸、电子地图,等等。

按信息载体形式也可简单地分成非电子资源和电子资源。

非电子信息源主要指的是靠人眼就能直接阅读的印刷型的传统图书、期刊等信息源。电子信息源主要是指电子图书、电子期刊、电子会议录等,也包括我们常说的缩微资料、视听资料、光盘资料等信息源。

1.2.2.3 按信息的出版形式分类

这是一种最常见的分类方法,包括图书、期刊、报纸、档案、标准、图谱、研究报告、会议资料、学位论文、专利说明书、产品说明书、政府出版物等。

(1) 图书(book)

图书大多是对已发表的科学技术成果、生产技术知识和经验经过著者的选择、鉴别、核对、组织而成的,论述比较系统,全面可靠,查阅方便(有目次表、索引),但出版周期较长,知识的新颖性不够。图书一般属于三次文献,但有的专著往往包含著者的新观点,或使用新的方法,新的材料,往往具有一次文献的意义。

图书种类较多,包括专著(monograph)和丛书(series of monographs)、教科书(textbook)、词典(dictionary)、手册(handbook)、百科全书(encyclopedia)等各种阅读型图书和参考书。从中可以看出,所谓的图书可以分为两种,一种为普通书籍,一种为工具书。

(2) 期刊(periodical)、报纸(newspaper)

期刊又称杂志(journal, magazine),一般是指具有固定题名,定期或不定期出版的连续出版物。其特点是出版周期短,报道文献速度快,内容新颖,发行及影响面广,能及时反映科学技术中新成果、新水平、新动向。期刊发表的论文大多数是原始文献,许多新的成果、新的观点、新的方法往往首先在期刊上刊登。科学技术的研究人员应熟悉本专业有关的期刊。常常阅读期刊可以了解动态,掌握进展,开阔思路,吸收新的成果。期刊论文是文献的主要类型,是检索工具报道的主要对象。

报纸是以刊载新闻和时事评论为主的定期向公众发行的连续印刷出版物。报纸是大众传媒的重要载体,具有反映和引导社会舆论的功能,对社会科学特别是对广泛的社会研究和企业经营来说,报纸是非常重要的信息源。我国第一份官方报纸是《北洋官报》(1902年,天津设立首局发行)。

(3) 报告(report)

报告是研究人员或企业围绕某一专题从事研究取得成果以后撰写的正式报告,或者是在研究过程中每一个阶段的进展情况的实际记录。其特点是内容详尽专深。报告的类型有技术报告(technical reports)、札记(notes)、论文(papers)、备忘录(memorandum)、通报(bulletin)、可行性报告(feasibility report)、市场预测报告(market prediction report)等。报告一般单独成册,有具体的篇名、机构名称和统一的连续编号(即报告号)。

报告一般划分为保密(classified)、解密(declassified)及公开(unclassified)等几种密级。保密的报告经过一定时间后往往会转为解密报告。非密资料中,又分为非密控制发行和非密公开发行。

(4) 会议文献(conference paper)

会议文献是指国际学术会议和各国国内重要学术会议上发表的论文和报告。此类文献一般都要经过学术机构严格的挑选,代表某学科领域的最新成就,反映该学科领域的最新水平和发展趋势。所以会议文献是了解国际及各国的科技水平、动态及发展趋势的重要情报来源。

会议的类型很多,归纳起来可分为国际会议、全国会议、地区性会议三种。会议文献大致可分为会前文献、会间文献和会后文献三类。

据统计,目前世界上每年有上万次学术会议,发表学术论文数十万篇。会议论文大都有新思想、新观点,是科学工作者所重视的情报源。

(5) 专利文献(patents)

专利是发明人创造发明了某种新技术,经政府专利局审批后,即获得一定年限的垄断权。专利权可以作为商品买卖。专利文献,主要是指专利说明书,分发明专利和实用新型专利、外观设计专利三种。这三种说明书是专利申请人向专利局递送的说明其发明创造的文件。在说明书中,发明人常常论述其发明解决了什么特殊问题,解决的方法,对旧有产品的改进及其他用途等。同时对企业引进技术和设备,以及保护企业自身利益的技术有着非常重要的作用。因此,专利文献已成了情报的一个重要来源。

(6) 学位论文(thesis 或 dissertation)

学位论文是高等学校、科研机构的研究生为获得学位,在进行科学研究后撰写的学术论文。学位论文一般要有全面的文献调查,比较详细地总结前人的工作和当前的研究水平,作出选题论证,并作系统的实验研究及理论分析,提出自己的观点。学位论文探讨的问题往往比较专一,带有创造性的研究成果,是一种重要的文献来源。

(7) 技术标准(technical standards)

技术标准是一种规范性的技术文件,它是在生产或科学研究活动中对产品、工程或其他技术项目的质量品种、检验方法及技术要求所作的统一规定,供人们遵守和使用。

技术标准按使用范围可分为:国际标准、区域性标准、国家标准、专业标准和企业标准等五大类型。每一种技术标准都有统一的代号和编号,独自构成一个体系。技术标准是生产活动中经常利用的一种情报信息源。

(8) 档案资料(archives, records, files)

档案是指具体工程、项目、产品和商品以及集团、企业的机构在技术和开发、运行、操作及活动过程中形成的文件、图纸、图片、方案、原始记录等资料。包括任务书、协议书、技术指标、审批文件以及研究计划、方案、大纲和技术措施,还包括相关的调查材料(原始记录、分析报告等)、设计计算、试验项目、方案、记录、数据和报告等,还包括设计图纸、工艺和其他相关材料。档案是企业生产建设和开发研究工作中用以积累经验,吸取教训和提高质量的重要文献,现在各单位都相当重视档案的立案和管理工作。

档案大多由各系统、各单位分散收藏,一般具有保密和内部使用的特点。它是各种社会活动的实录,是真实可靠的历史信息情报,具有很高的参考价值。

(9) 政府出版物(government publication)

政府出版物是各国政府部门及其所属的专门机构发表、出版的文件,其内容广泛,从基础科学、应用科学到政治、经济等社会科学无所不包。就文献的性质来看,其内容可分为行

政性文件(如政府法令、法规、方针政策、调查统计资料等等)和科技文献(科技报告、科普资料、技术政策等)两大类。通过这类文献可了解一个国家的科学技术、经济政策、法令、规章制度等。这类资料具有极高的权威性,对企业的活动具有重要的指导性。

(10) 产品样本(catalogue)

产品样本是国内外生产厂商或经销商为推销产品而印发的企业出版物。用来介绍产品的品种、特点、性能、结构、原理、用途和维修方法、价格等。查阅、分析产品样本,有助于了解产品的水平、现状和发展动向,获得有关设计、制造、使用中所需的数据和方法,对于产品的选购、设计、制造、使用等有着较大的参考价值。

由于产品样本是已经生产的产品说明,在技术上比较成熟,数据比较可靠,对产品的具体结构、使用方法、操作规程、产品规格都有较具体的说明,并常常附有外观照片和结构图。专利产品还注有专利号(根据专利号可查找专利说明书),对于新产品的的设计、试制都有较大的实际参考价值。

1.2.2.4 按信息的加工深度分类

按信息的加工程度对信息进行划分,信息可分为一次信息、二次信息、三次信息。

一次信息也称原始信息,是指人们根据自己的科学实验、生产实践的成果而撰写的文献,一般指公开出版的图书、期刊论文、科技报告、会议论文、学位论文、专利等。无论信息存储于何种载体,只要是原始资料就是一次信息。

二次信息也称检索工具,是指对一次信息加工、整理后形成的各种检索工具,如目录、题录、文摘等,它不对一次信息提供评论,仅提供一次信息的检索线索,也就是说,二次信息是把一次信息从分散、无序状态变成集中、有序化的状态,供人们迅速、准确地查询所需信息。

三次信息是对一、二次信息按知识门类或专题进行综合加工的产物,如百科全书、年鉴、手册、综述等。

除此之外,还有零次信息(文献),指的是信息的内容尚未经专门机构加工整理就直接作用于人的感觉信息情报,是未经发表和有意识处理的最原始的资料。零次信息主要包括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、未公开于社会即未经正式发表的原始的文献,如书信、手稿、记录、笔记等。零次信息一般通过口头交谈、参观展览、参加报告会等途径获取。

第2章 信息检索基础

2.1 信息检索原理

2.1.1 信息检索的含义

信息检索(information retrieval)是指将信息按一定的方式组织和贮存起来,并根据信息用户的需要找出有关信息的过程。所以,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。狭义的信息检索则仅指该过程的后半部分,即根据课题的需要,主要借助于检索工具,从信息集合中找出所需信息的过程,相当于人们所说的信息查找(information search)。信息检索的过程往往需要一个评价反馈途径,多次比较匹配,以获得最终的检索结果。从而可知,信息检索的全过程应包括以下两个方面的内容:

(1) 信息标引和存储过程

是指对大量无序的信息资源进行标引处理,使之有序化,按科学的方法存储,组成检索工具或检索文档,即组织检索系统的过程。

(2) 信息的需求分析和检索过程

是指分析用户的信息需求,利用已组织好的检索系统,按照系统提供的方法与途径检索有关信息,即检索系统的应用过程。因此,信息检索是将描述特定用户所需信息的提问特征,与信息存储的检索标志进行异同的比较,从中找出与提问特征一致或基本一致的信息,并将结果显示给用户。提问特征是对信息的需求进行分析,从中选择出能代表信息需求的主题词、分类号或其他符号。检索标志是指通过对查新项目主题分析将自然语言转换成规范化语言,即确定检索入口的问题。检索时,将提问特征与检索标示进行对比匹配,检索结果如显示一致或基本一致,即为所需信息。

2.1.2 信息检索的作用

由于信息技术的发展,信息资源的激增,从而使得学习信息检索知识和技能成为人们知识结构中不可缺少的重要组成部分。这对于提高人们的信息素养,培养复合型、开拓型人才具有很明显的作用。

(1) 信息检索是获取知识的捷径

在我们的工作、学习、研究中,经常会碰到这样那样的问题,要在浩瀚的信息资源海洋里一索即得,解疑释惑,仅仅只有一定的基础知识和专业知识还远远不够。我们还必须学习信息检索的方法,从而找到一条吸收和利用大量新知识的捷径,把自己带到更广阔的知识领域中去,不断拓宽知识面,调整知识结构,对未知世界进行探索,才能游刃有余地解决所遇到的问题。

(2) 信息检索是科学研究的导向

科学技术的迅猛发展加速了信息的增长,加重了信息用户搜集信息的负担。任何人从事某一特定领域的学术活动,或开始做一项新的研究工作,都要花费大量时间,从查找资料,寻求信息开始,到对有关研究工作全面的调查研究、分析、评价、预测,充分了解国内外是否有人做过或者正在做同样的工作,了解取得的成果、存在的问题等都是必须要花时间和精力去解决的。为了提高效率,提高文献的检全率和检准率,信息检索就是研究工作的基础和必要环节。科研工作者不仅要具有实际的科学研究能力,而且还应具有信息收集、选择和利用的能力。成功的信息检索无疑会节省研究人员的大量时间,使其能用更多的时间和精力进行科学研究,达到事半功倍的效果。

(3) 信息检索是终身教育的基础

学校培养学生的目标是学生的智能:包括自学能力、研究能力、思维能力、表达能力和组织管理能力。联合国教科文组织(UNESCO)提出,教育已扩大到一个人的整个一生,认为唯有全面的终身教育才能够培养完善的人,不断更新知识,可以防止知识老化,适应当代信息社会发展的需求。所以,学习信息检索,不但要学会查找信息资源的方法,更重要的是要通过本课程的学习,提高自身的信息意识和信息观念,提高独立分析问题和解决问题的能力,使自己具有更强的社会生存和社会竞争能力,终身受益。

2.1.3 信息检索原理

信息检索包括信息的存储和检索两个过程:是将信息概念进行分析,通过标引(分类、主题、题名、作者、代码、字段等),依据标引组成信息系统;信息用户则依据标引在信息系统中检索有关信息。

信息系统指的是有组织的信息集合。就是将大量的信息,用具体的数据项标志、标引、排列和存储起来。它可以是图书馆、信息中心的全部馆藏、数据库的全部记录。

文献检索的核心问题是实现文献特征与用户提问需求特征之间的匹配,为此,用户一方面要掌握文献集合的组织编排方式;另一方面也要善于分析个人的信息需求,善于将信息需求转化为可检索的概念术语。这是一个双向分析和匹配的过程,是一个以满足用户信息需求为目标,以用户提问为线索,不断缩小检索范围并查找用户所需文献的过程。

由此可知,信息检索的本质就是读者(用户)的信息需求与存储在信息集合体中的信息进行比较和选择,即匹配的过程。也就是对一定的信息集合体(系统)采用一定的技术手段,根据一定的线索与准则找出相关的信息。存储是为了检索,没有存储就无所谓检索。信息的存储与检索相辅相成,相互依存。

2.1.4 信息检索类型

信息检索根据检索的目的和对象不同,可以分为书目信息检索、全文信息检索、数据信息检索和事实信息检索。

(1) 书目信息检索

以标题、作者、摘要、来源出处、专利号、收藏处所等为检索的目的和对象,检索的结果与课题相关的一系列书目信息线索,即检索结果不直接解答课题用户提出的技术问题本身,只提供与之相关的线索供参考,用户通过阅读后才决定取舍。因此,书目信息检索是一种相关性检索。