

GongCheng XinXi

信息检索丛书

JianSuo

工程信息检索

李爱国 主编

东南大学出版社

信息检索丛书

工程信息检索

主 编	李 爱 国			
副 主 编	潘 松 华	杨 新 村		
主 审	张 智 松			
编 者	李 爱 国	潘 松 华	杨 新 村	
	陆 美	朱 晓 慧	胡 彬	
	魏 彬	朱 佳 鸣	王 舞 阳	
	张 明 果	隆 新 雯	马 青	
	王 晓 梅	李 晓 鹏	刘 利	
丛书主编	滕 跃	蔡 忆 宁		

东南大学出版社

• 南 京 •

内 容 提 要

本书主要介绍信息检索基础知识、工程信息检索平台、工程信息检索系统、专利信息检索、搜索引擎、主题网关检索、科技创新和学术规范等内容,详细介绍了检索途径及检索方法。

本教材内容新颖,学科重点突出,适合本科生、研究生作为文献检索教材,同时也可作为教师及工程领域的科技研究人员文献检索的指导工具。

图书在版编目(CIP)数据

工程信息检索/李爱国主编. —南京:东南大学出版社,2005.9

ISBN 7-5641-0147-4

I. 工... II. 李... 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 102894 号

出版发行:东南大学出版社

社 址:江苏省南京市四牌楼 2 号(210096)

出 版 人:宋增民

经 销:江苏省新华书店

印 刷:南京京新印刷厂

版 次:2005 年 9 月第 1 版 2005 年 9 月第 1 次印刷

开 本:700mm×1000mm 1/16

印 张:21.5

字 数:433 千

印 数:1~10 000 册

定 价:29.00 元

(若有印装质量问题,请同读者服务部联系。电话:025-83792328)

序

以东南大学图书馆李爱国博士为主编,盐城工学院图书馆潘松华馆长、华东电力实验研究院科学技术信息研究所杨新村所长为副主编的《工程信息检索》一书的问世,是一件值得庆贺的事。它的诞生将成为当前百花盛开,群芳争妍文献检索课教科书园地中的一朵奇葩,以它特有的编撰体例和深入浅出的编写方法让人刮目相看。

自国家教育部 1984 年 2 月下发的《关于在高等学校开设〈文献检索与利用〉课的意见》以来,全国高校陆续地开设了文献检索与利用课,并编写了一大批系列教材。有综合性的,亦有专科性的;有社会科学的,亦有自然科学与工程技术的;同时亦出有单体的专门化资料。凡此种种为当时提升在校的高校学生在资料查检的能力方面起到了积极的效果,对他们信息素质的提高提供了极大的帮助。

20 世纪 90 年代以来,以通讯技术、网络技术、计算机技术相结合的现代技术的问世,以 Internet 为核心连接起来的全球计算机网络,使传统的相对集中和规范为基础的文献数据及其检索系统面临严峻的挑战。具体表现为:以缩微品、声像带、磁盘、光盘等形式记录的非纸形信息急剧增加;纸质文献与电子文献并存;内容涉及各学科领域乃至人类生活的方方面面;各类型、各品牌的网络信息查询工具如雨后春笋般的不断涌现;信息资源数据库分布在世界各地主机上,网络化与数字化技术又将其连为一体,为人们跨时空、高效快速的传递信息提供了国际化的知识信息检索平台,极大地提高了人们获取信息的可能性和有效性。诚然,面对社会的这一态势,靠“手翻、眼看、大脑判断”的手工检索方式就显得苍白无力,难以全面适应当今信息社会发展的需要。与此同时,又有不少专家在其原教科书的基础上增加了计算机检索、网络检索、多媒体检索等内容。就当时而言,这一举动对信息检索学科的发展起到了推进的作用,收到了良好的效果。但就成书体例来讲,仍没有脱开原有的框框,仍然是分别介绍各种检索工具书的编排体例和查找途径等内容。至于计算机检索只不过是对其技术的阐述,最好的也只是介绍了 DLALOG 系统检索方法而已。鉴于诸如此类问题的存在,李爱国、潘松华等同志在总结教学经验的基础上,根据社会需求的实际状况,编撰了与众不同的《工程信息检索》一书,脱开了同类图书编撰的体例和编写方法,起到了“授之以渔”的效果。

该书共由八章构成。与其他同类图书不同之处是:全书完全是建立在现代信息检索技术基础之上而成型的。开始,在强化了信息素养、信息源及信息检索语言等必

要的信息检索基础知识之后,又以信息检索平台形式介绍了7个中外著名的二次文献数据库;以信息检索系统的方式介绍了11个中外著名的全文数据库。随后,又介绍了工程类须臾不能离开的专利和标准信息。这样编写的方式,不仅体现了该书的系统性、严谨性和科学性,而且不会使人对数据库有割裂之感觉,让人明白不同的数据库所能帮助你解决的问题。网络信息检索已成为当今世界人们查检各种信息的一种主流。但是,如何准确、有效地查到自己所必需的知识信息,那就是要对搜索引擎和主题网关的知识有一个详细的了解。本书在这两方面均给予了详尽的阐述。阅读以后,就会使你能够正确的、有效的利用搜索引擎和网关的方法知识去查找到分散在各地的网上信息。也许这是本书的重大突破。信息检索的目的在于科技创新。作者在该书中将此内容另立一章,阐明了国家创新体系、信息新颖性查证和项目技术要点及查新理论等内容,促使人们对信息新颖性的鉴别能力得到大大的提高。当今网络信息多如牛毛,给人们甄别信息的真伪带来了一定的难度。作为信息检索课教材,加强学术规范意识的内容,提高信息识别的能力是非常必要的。因此,该书作者在最后章节阐明了学术规范的基本原则和法律法规,以及学术的评价规范、批评规范和引用规范等内容是很有意义的,这在当前同类教科书中尚属首例。

总之,本书观点新颖,资料翔实,系统性、严谨性和科学性是该书保证高质量的必要措施,这也充分说明了作者们的良苦用心,他们的责任意识、大局意识和精品意识是应该加以肯定的。我相信,每一位通过阅读此书、理解此书的人,都有可能成为信息检索的行家里手,在网络上能够游刃有余地检索出每一条信息。

李爱国、潘松华等同志及其合作者们,长期以来一直从事图书馆信息的开发与服务、信息资料查新及研究等工作,在管理、教学和图书馆信息化方面颇有建树。20世纪90年代初,他们就从事了科技信息查新、信息检索和教学等工作,还直接参与了东南大学图书馆 DILAGO 国际联机情报检索终端的建设工作,并出版过一些专著,同时,他们又都有高级技术职称,因此由他们来编撰该书是值得信赖的。我认为该书除作为教材使用外,还可以作为从事工程技术的科技工作者和广大青年读者们自学读物,同时也适用于其他专业人员使用,可以说它将成为我们的案头工具。我想,该书的问世,对信息资源的开发利用,对当前学生的素质教育和创新能力的培养,以及对我国国民经济的建设将发挥积极的作用。当然,由于时间紧,任务重,书出众手,且带有一定的探索性,书中的缺点与不足之处在所难免,尚需从事信息教学和实践的同仁们匡正。经过广泛听取意见,使这部教材日臻完善,在未来的教学科研中发挥出它应有的作用。

张智松

2005年8月9日

于南京师范大学图书馆

前 言

随着科学技术的不断进步和新的信息技术层出不穷,人们的信息意识、信息观念以及获得信息的手段也在不断地更新变化。

信息检索的意识和能力是考察一个人信息素养的重要指标。可以说,提高信息检索的意识和能力是每个人必修之功课。本书共分信息检索基础知识、工程信息检索平台、工程信息检索系统、专利信息检索、搜索引擎、主题网关检索、科技创新和学术规范等八章,附录中列举一些通过普通搜索引擎不可搜索的“深网”(Deep Web)站点。本书具有以下特点:首先,概述信息检索的基础知识,综合各工程信息检索系统共性,逐步建立信息检索是“以人中心”而非“以计算机系统为中心”的意识;其次,搜索引擎是重要的网络信息检索工具,但是其对信息源的组织、加工非人工化或较少的人工干预影响着信息检索的结果,所以本书将主题网关检索纳入其中,在类似的书本中是首次尝试;最后,本书将科技创新作为单独章节进行论述,目的是向读者说明文献学对科学学有着重要的支持作用,鼓励大家遵守学术规范,进行科技创新活动。

本书的编写得到东南大学图书馆、南京师范大学图书馆、盐城工学院图书馆、上海华东电力实验研究院科技信息所、江苏省科学技术情报研究所等单位的支持,在此表示衷心的感谢!

由于时间仓促,加之编者水平限制,不足之处请读者与同行不吝赐教。

李爱国

2005年8月于南京

目 录

1 信息检索基础	1
1.1 信息及信息检索与信息素养	1
1.1.1 信息及其相关术语	1
1.1.2 信息检索与信息素养	5
1.2 信息源	8
1.2.1 信息源与信息资源	8
1.2.2 文献信息源载体类型	10
1.2.3 文献信息源加工类型	11
1.2.4 文献的出版类型及著录特征识别	16
1.2.5 电子信息源	19
1.3 信息检索理论与方法	22
1.3.1 信息检索原理	22
1.3.2 信息检索类型	23
1.3.3 检索手段及技术	25
1.3.4 信息检索途径及方法	29
1.3.5 信息检索步骤	32
1.3.6 信息检索效果评价	33
1.4 信息检索语言	35
1.4.1 信息检索语言的概念	35
1.4.2 信息检索语言的类型	35
1.4.3 分类检索语言	37
1.4.4 主题检索语言	38
2 工程信息检索平台	41
2.1 Engineering Village 2	41
2.1.1 概况	41
2.1.2 检索方法	42
2.1.3 检索步骤	42

2.1.4 检索技巧	43
2.2 ISI Web of Knowledge	53
2.2.1 Thomson 公司和 ISI 概述	53
2.2.2 ISI Web of Knowledge 概况	54
2.2.3 ISI Web of Knowledge 提供的服务	56
2.2.4 Web of Science	56
2.2.5 Web of Science 的收录范围	57
2.2.6 Web of Science 的检索方法	57
2.3 OVID 检索平台	65
2.3.1 OVID 概况	65
2.3.2 BIOSIS Preview 数据库的检索	66
2.4 《剑桥科学文摘》数据库(CSA)	69
2.4.1 概况	69
2.4.2 CSA 检索系统	70
2.5 MathSciNet(数学评论)	73
2.5.1 概况	73
2.5.2 MathSciNet 数据库检索系统	73
2.6 ProQuest-PQDD	76
2.6.1 概述	76
2.6.2 PQDD 的收录范围	77
2.6.3 检索方法	78
2.6.4 ProQuest 学位论文全文数据库的检索方法	83
2.7 万方数据资源系统	86
2.7.1 概况	86
2.7.2 万方数据资源系统的检索	86
3 工程信息检索系统	94
3.1 Elsevier EJOS	95
3.1.1 Elsevier EJOS 概况	95
3.1.2 Elsevier EJOS 的检索方法	96
3.1.3 Elsevier EJOS 的检索技术	99
3.1.4 Elsevier EJOS 的检索结果	100
3.2 SpringerLink	102
3.2.1 SpringLink 概况	102
3.2.2 SpringerLink 的检索方法	103

3.2.3 SpringerLink 的检索技术	105
3.2.4 SpringerLink 的检索结果	106
3.2.5 SpringerLink 的个性化服务	109
3.3 IEL	111
3.3.1 IEL 概况	111
3.3.2 IEL 的检索方法	112
3.3.3 IEL 的检索结果处理	116
3.3.4 IEL 的个性化服务功能	117
3.4 Wiley InterScience	118
3.4.1 Wiley InterScience 概况	118
3.4.2 Wiley InterScience 的检索方法	119
3.4.3 Wiley InterScience 检索技术	121
3.4.4 Wiley InterScience 检索结果	121
3.4.5 Wiley InterScience 的个性化服务	122
3.5 ACM Digital Library	124
3.5.1 ACM 概况	124
3.5.2 ACM 的检索方法	124
3.5.3 ACM 的检索技术	126
3.5.4 ACM 检索结果处理	127
3.6 ACS e-Journals	128
3.6.1 ACS e-Journals 概况	128
3.6.2 ACS e-Journals 检索方法	129
3.6.3 ACS e-Journals 的检索技术	131
3.6.4 ACS e-Journals 的检索结果处理	131
3.6.5 ACS e-Journals 的个性化功能	132
3.7 IOP	133
3.7.1 IOP 概况	133
3.7.2 IOP 的检索方法	133
3.7.3 IOP 检索结果处理	136
3.7.4 IOP 个性化服务功能	136
3.8 Scitation 平台	138
3.8.1 概述	138
3.8.2 Scitation 平台的使用	141
3.9 EBSCO 数据库	146
3.9.1 概况	146

3.9.2 检索方法	147
3.10 清华同方系列数据库(CNKI)	153
3.10.1 CNKI 概况	153
3.10.2 “中国期刊网”全文数据库检索系统	154
3.11 中文科技期刊数据库(VIP)	159
3.11.1 重庆维普资讯有限公司概况	159
3.11.2 中文科技期刊全文数据库	161
4 专利信息检索	167
4.1 专利和专利文献	167
4.1.1 专利文献的特点	167
4.1.2 专利文献的作用	168
4.1.3 专利文献的体裁格式	169
4.1.4 专利分类	169
4.1.5 印刷型专利文献的检索	172
4.1.6 网络专利数据库检索	172
4.1.7 国内外网络专利数据库简介	173
4.2 德温特创新索引 DII	175
4.2.1 德温特的 WPI 体系	176
4.2.2 德温特专利引文索引(PCI)	177
4.2.3 DII 数据库概述	178
4.2.4 DII 检索方法与步骤	179
4.2.5 DII 检索结果	184
4.3 USPTO	186
4.3.1 美国专利文献概述	186
4.3.2 美国印刷型专利检索工具	187
4.3.3 USPTO 专利数据库	189
4.4 欧洲专利检索	193
4.4.1 esp@cenet 概述	193
4.4.2 esp@cenet 检索	194
4.5 日本专利检索	197
4.5.1 日本专利及专利文献概述	197
4.5.2 日本专利数据库(IPDL)	199
4.6 中国专利文献及其检索	204
4.6.1 我国印刷型专利文献的检索工具	204

4.6.2 中国专利文献网络数据库	206
5 搜索引擎	214
5.1 搜索引擎概述	214
5.1.1 搜索引擎的产生与发展	214
5.1.2 搜索引擎的原理	215
5.1.3 搜索引擎的分类	216
5.1.4 搜索引擎结果的排序技术	218
5.2 搜索引擎使用	221
5.2.1 中文综合性搜索引擎	221
5.2.2 外文综合性搜索引擎	236
6 主题网关检索	254
6.1 概述	254
6.2 主题网关的分类	257
6.3 “主题网关”检索举要	257
6.3.1 Aecportico	257
6.3.2 Biogate	260
6.3.3 EEVL	264
6.3.4 The Math Forum Internet Mathematics Library; mathematics	268
6.3.5 PSIGate	271
6.3.6 ChemDex; chemistry	274
6.4 Renardus 服务	276
7 科技创新	279
7.1 科技文献与科技创新	279
7.1.1 文献信息资源是知识创新的基础与源泉	279
7.1.2 科技信息资源及其服务体系的构建	280
7.1.3 科技情报与情报学	285
7.2 科技创新体系	286
7.2.1 创新	287
7.2.2 国家创新体系	288
7.3 科技查新	290
7.3.1 科技查新咨询工作的性质与要求	290

7.3.2 新颖性查证的针对性与分析深度是查新工作的立足之本	291
7.3.3 新颖性的概念与内涵	292
7.3.4 项目技术要点与查新点	295
7.3.5 新颖性判断	297
8 学术规范	303
8.1 学术规范的基本概念	303
8.2 学术活动基本法律法规	304
8.3 学术道德规范基本原则	306
8.4 学术论文写作规范	308
8.5 学术评价规范	313
8.6 学术批评规范	316
8.7 学术引用规范	317
8.8 加强学术规范教育的意义	318
附录	319
1 部分标准文献搜索站点	319
2 部分 Deep Web 搜索引擎	320
3 部分 Deep Web 数据库	321
4 部分知名搜索引擎	322
5 部分主题网关	323
参考文献	327

1

信息检索基础

1.1 信息及信息检索与信息素养

信息无时无刻不在,与人类的发展历史更是紧密相系。不管是日常生活还是科学研究,人们都在自觉或不自觉地利用信息。理解信息的涵义,知道信息资源,掌握信息获取技巧,懂得如何利用信息服务是今天的大学生必须拥有的最基本的知识之一。

对于工程技术人员、研究人员来说,首先应掌握工程技术信息。因为工程技术具有很强的继承性,积累并继承了前人和他人的成果,是技术进步和发展的重要前提。工程技术文献是工程技术进步和发展的记载,因此,通过信息检索,才能及时掌握和运用与本专业有关的技术信息,才能进行技术再创造,才能推动技术不断创新向前发展。

1.1.1 信息及其相关术语

(1) 信息的涵义与类型

① 信息的涵义

人们对信息的利用十分广泛,从而对信息的研究也不断深入。截至今日,人们对信息的认识和理解十分多样化。据不完全统计,信息的定义有 100 多种,至今仍没有统一的、能为各界普遍认同的定义。

最早提出信息(Information)这一概念的是 C. E. Shannon 申农(香农)和 W. Weaver(维纳),他们试图给信息一个正式的和定量的定义。

信息论的创始人申农(香农)从通信系统理论的角度把信息定义为:用来减少随机不确定性的东西(这个观点是情报学界引述最多的)。他的意思即信宿(信息接受方)未收到消息前不知道信源(信息产生方)发出什么信息,只有在收到消息后才能消除信源的不确定性。如果没有干扰,信宿得到的信息量与信源的不确定性相等。他的看法被认为是对信息认识的重大进展,因为他推导出了信息测度数学公式,标志着

信息科学进入了定量研究阶段。

控制论专家维纳从信息发送、传输、接收的过程中,客体和接收主体之间的相互作用做出定义。他说,“信息既不是物质,又不是能量,信息就是信息。”

从经济学的角度看,信息是与物质、能量相并列的客观世界的三大要素之一,是为管理和决策提供依据的有效数据。

从认识论的角度看,信息既是客观存在的,又是人的主观认识的产物,它不同于客观世界,也不同于精神世界,而是物质与精神的中介。

从图书馆学的角度看,信息被定义为事物或记录,它们可以是文本、图片、录音磁带、博物馆陈列品、自然物体、实验、事件等。

我国《辞海》对信息的释义是:音讯、消息;通讯系统传输和处理的对象,泛指消息和信号的具体内容和意义。

美国《韦氏字典》对信息释义为:接受由他人传来的知识,这些知识是由他们通过调研和学习获得的;从特定的事件或状况下,如消息、新闻、通知中获得的知识;用来传播的事实和数据。

社会的进步赋予信息更丰厚的内涵,导致人们对信息的认识更加深入而众多,信息的膨胀与人们对其需求的激增,又使信息成为当今社会生活的一大支柱,成为一种与能源、材料并存的重要战略资源。

② 信息的特点

客观性 信息是客观存在的,它源于客观存在的物质及其运动两大特性。信息的客观性还表现为:它是以物质的客观存在为前提的,即使是主观信息,如决策、判断、指令、计划等,也有其客观实际背景,并以客观信息为“原料”,受客观实践检验的。

依附性 信息总是依附于一定的物质载体而存在的,需要某种物质承担者。声音、语言、文字、颜色、图像、生物等各种信息信号、纸张、胶片、人的大脑等,无一不是信息的载体。

可传递性 信息的产生是同信息的传递联系在一起的,信息在传递过程中发挥作用。信息的传递过程是一个重复使用的流通过程,在这一过程中,信息的占有者不会因传递信息而失掉信息,一般说来,也不会因多次使用而改变信息的自身价值。

可塑性 信息可以加工处理,可以压缩、扩充和叠加,也可以变换形态。

时效性 现代社会中,信息的使用周期迅速缩短,信息的价值实现取决于及时地把握和运用恰当或正确的信息,如天气信息、经济信息、交易信息、科学信息等。

共享性 信息的共享性主要表现在同一内容的信息可以在同一时间内由两个或两个以上的使用者同时使用,且不会失去所提供的信息内容和信息量,这一点是信息资源的广泛提供与利用的基础。

③ 信息的类型

信息广泛存在于自然界、人类社会和思维领域。可以这么说:自然界与人类活动

的事实及人类对它们的认识和创造是信息的内容组成,而载体记录和媒体传播则是信息存在的物理形式。

从产生信息的客体的性质来分,信息可分为自然信息、生物信息和社会信息。

自然信息 自然信息包括瞬时发生的声、光、热、电及形形色色的天气变化、缓慢的地壳运动、天体演化,以及作用在树干上的年轮等。

生物信息 生物信息是生物为繁衍生存而表现出来的各种形态和行为,如遗传信息、生物体内信息交流、动物种群内的信息交流。

社会信息 社会信息是指人与人之间交流的信息,既包括用手势、身体、眼神所传达的非语义信息,也包括用语言、文字、图表等语义信息所传达的一切对人类社会运动变化状态的描述。按照人类活动领域,社会信息包括科技信息、政治信息、军事信息、经济信息和文化信息。

以信息所依附的载体为依据,可分为文献信息、口头信息、电子信息等。

文献信息 文献信息就是文献所表达的内在信息,以文字、符号、声像信息为编码的人类精神信息,也是经人们筛选、归纳和整理后记录下来的信息。

口头信息 口头信息指存在于人脑记忆中、通过交谈、讨论、报告等方式交流传播的信息。口头信息具有出现早、传递快、偶发性强的特点,但缺乏完整性和系统性,大部分转瞬即逝。

电子信息 电子信息是计算机技术、通信技术、多媒体技术和高密度存储技术迅速发展的产物,是当今发展最快、最具应用价值和发展前途的新型信息源。

此外,信息分类还有其他划分方法,如以信息的记录符号为依据,可分为语音信息、图像信息、文字信息、数据信息等;以信息的运动状态为依据,可分为连续信息、离散信息;以信息的加工层次而论,可分为原始信息和再生信息(二次信息、三次信息)。

(2) 知识

知识(Knowledge)是人类通过信息对自然界、人类社会及思维方式与运动规律的认识与概括,是人的大脑通过思维重新组合的系统化了的信息,是信息中最有价值的部分。知识可分为理性知识和感性知识。理性知识(主观知识、隐性知识),是对客观事物的本质和规律性的认识,是经过思维、逻辑加工的知识,构成知识体系;感性知识(客观知识、显性知识),是对客观事物的描述和对现象、事实的感知,是未经逻辑加工的知识。

联合国研究机构 1990 年提出“知识经济”一词,经济合作与发展组织(Organization for Economic Cooperation and Development, 简称 OECD)1996 年明确提出:知识经济是建立在知识和信息的生产、分配和使用之上的经济,以创造性的人力资源为依托,以高科技产业及智业(信息咨询、管理服务)为支柱的经济,并出版了《以知识为基础的经济》报告,提出知识又可分为四种类型,即 4W:

知事(know-what)——掌握事实;

知因(know-why)—理解自然原理和规律方面的科学理论;
 知道怎样做(know-how)—做事的技能,尤其是内部的技术诀窍或专有技术;
 知道谁(know-who)—确定谁拥有信息和专长,
 包括其创造思想、方法、手段、过程。

人们习惯把“知事”看作信息,但其实上述四种类型的知识都与信息之间存在着密切的关系。信息与知识的关系可以用图 1-1 来表示。

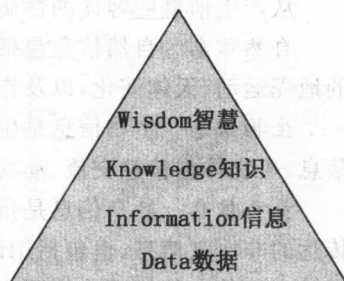


图 1-1 信息与知识的关系

从图 1-1 同时可以看出信息的结构以及信息与知识的关系。信息依赖于数据,众多的数据集成信息,通过对信息的系统化的认识转化为知识,而足够的知识则可以提高智慧。这四个元素之间呈现出“金字塔”型的关系。在这个“金字塔”中,信息是知识的重要组成部分。但是,只有将信息经过加工,成为对自然和社会发展客观规律的认识时,这种再生的信息才构成知识。英国著名情报学家 B. C. 布鲁克斯这样表述信息与知识的关系:“信息是使人原有的知识结构发生变化的那一小部分知识”。

人们系统化认识信息并转变成知识有两个途径: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的传递、交流而间接获得。然而,获得的信息能否转化为知识,转化得是否充分、完整,则取决于受主的认知能力。

在知识经济时代,不管是科研、管理还是工程技术领域各行各业的人们都感受到更新自身的知识结构的需要,所有这些信息需求在一定的条件下都将激发获取信息的动机并最终转化为查找信息的行为。

(3) 文献

我国《国家标准文献著录总则》对文献(Literature)有如下定义:文献,即记录有知识的一切载体。

《文献情报术语国际标准(草案)》(ISO/DIS5127)对“文献”的定义是:为了把人类知识传播开来和继承下去,人们用文字、图形、符号、声频、视频等手段将其记录下来,或写在纸上,或晒在蓝图上,或摄制在感光片上,或录到唱片上,或存储在磁盘上。这种附着在各种载体上的记录统称为文献。

从上述定义可以看出现代文献由四要素构成:文献信息、文献载体、符号系统和记录方式。四要素缺乏任何一种都不可能形成文献。文献信息是文献的内容,它属于信息,具有信息的一切性质和功能;符号系统是信息的携带者,包括图画、文字、公式、图表、编码、声像和电磁信息等;记录方式是代表文献的符号通过特定的人工记录方式进入载体的方法和过程;载体是符号的依赖物,即材质,它在人类文明的演进过程中一直在不断地演变:从石板、兽骨、树叶到陶、玻璃、青铜、金银、纸、胶片、磁带、光盘等。目前,除了纸质(印刷型)文献外,常见的其他介质的文献主要有以下类型:缩

微文献,声像型文献,电子型文献(磁盘和光盘)。

文献对人类文明进步具有重大意义,主要起着两方面的作用,其一是知识的积累、传播与创造,其二是对研究成果的确认与评价。研究人员充分利用他人的研究成果、他人的文献信息,通过自身的研究活动去产生新的知识,总结成新的文献,再供其他人享用。这样,形成一个知识、文献信息不断充实、增长的信息流和知识链。人们在推动世界进步的同时也实现了个人的价值。

1.1.2 信息检索与信息素养

信息具有共享性。信息资源共享(Information Resource Sharing)是当今社会的一个热点问题。为了分享人类共同的知识财富、研究成果,人们必须通过一种科学的方法从取之不尽的信息源中去识别和获取所需要的那部分信息,这个过程就是检索(Searching)。

随着信息加工技术的发展,人类创生的信息量也在高速增长。据统计,每年出版图书 50 万种以上、期刊 10 万种以上,每年发表的科技论文总数近 500 万篇,并呈指数级增长。信息浩如烟海,如何从这浩如烟海的信息中找出所需信息,就成为信息检索的重任。

(1) 信息检索的定义

信息检索(Information Retrieval),是指将信息按一定的方式组织和存储起来,并根据信息用户的需要找出相关信息的过程,所以它的全称又叫“信息的存储与检索”(Information Storage and Retrieval),信息的存储方式与检索方式是对应的,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,即从信息集合中找出所需要的信息的过程,相当于人们通常所说的信息查寻(Information Search)。

当前更普遍使用的术语是信息存取(Information Access),该词源于计算机科学领域,指计算机访问文档或数据集的方式。引入信息检索范畴,则从本质上拓宽了检索的内涵及其应用。信息存取将所有信息的组织、检索活动及其先进的技术手段融合在一起。如果检索的对象是文献信息,那么就属于文献检索。

不管是信息检索,还是信息存取,均可以增强获取新知识的能力,可以用来构造、补充或完善自己的知识结构,人们将其称之为信息技能。

(2) 信息检索与创新

信息具有可扩充性,人们对信息的处理能力越强,信息扩充得就越快,即知识创新就越快、增值也越快。检索者能从当今庞大而无序的信息海洋中准确快速地找到所需的信息,这就是信息检索能力,这种能力恰恰就是其创新能力的重要组成部分,在信息处理和增值中有着重要的意义。

信息的检索、利用和创造是一个循环和增值的过程,人们通过工具的检索获得信息,经过处理筛选出需要的部分,在利用信息的过程中又创出新的信息,这些信息经