

全国高等教育精品规划教材

信息检索教程

XINXI JIANSUO JIAOCHENG

主 编 王 宇



高等教育出版社
HIGHER EDUCATION PRESS

全国高等教育精品规划教材

信息检索教程

XINXI JIANSUO JIAOCHENG

主 编 王 宇

副主编 吴 瑾 刘允斌 宫 平



高等教育出版社·北京

HIGHER EDUCATION PRESS BEIJING

图书在版编目(CIP)数据

信息检索教程 / 王宇主编. —北京: 高等教育出版社, 2013. 3 (2015. 1 重印)
ISBN 978-7-04-037075-1

I. ①信… II. ①王… III. ①情报检索 - 高等学校 - 教材 IV. ①G252.7

中国版本图书馆 CIP 数据核字 (2013) 第 048756 号

策划编辑 肖春茂 责任编辑 肖春茂 封面设计 吴昊 责任印制 蔡敏燕

出版发行	高等教育出版社	咨询电话	400-810-0598
社 址	北京市西城区德外大街 4 号	网 址	http://www.hep.edu.cn
邮政编码	100120		http://www.hep.com.cn
印 刷	上海华教印务有限公司		http://www.hepsh.com
开 本	787mm×960mm 1/16	网上订购	http://www.landracom.com
印 张	16.75		http://www.landracom.com.cn
字 数	320 千字	版 次	2013 年 3 月第 1 版
购书热线	010-58581118	印 次	2015 年 1 月第 4 次印刷
	021-56717287	定 价	33.00 元

本书如有缺页、倒页、脱页等质量问题, 请到所购图书销售部门联系调换

版权所有 侵权必究

物 料 号 37075-00

前言

随着知识经济和科学技术的迅猛发展,信息因载体多元化而表现出形式多样化和处理复杂化,获取、评价和使用信息的能力显得更为重要,信息素养也成为评价人才综合素质的一项重要指标。信息素养教育是高等教育中必不可少的重要部分,其目的是培养大学生的信息意识、获取信息的能力,从而提高其创新能力。

信息素养教育的主要形式是信息检索课,其主要任务是让大学生掌握获取信息的科学方法,了解专业信息检索的基本知识,学会常用检索工具的使用方法,能够综合利用各类信息资源,从而增强自学能力和研究能力。沈阳师范大学于2003年开设“信息检索与利用”课程,授课对象覆盖全校各学科的大二至大四的本科生及研究生。2007年成立了信息检索教研室。多年来,信息检索教研室明确教学目标,创新教学理念,改革教学方法和教学内容,搭建学习平台,加强实践力度,切实提高了学校学生的信息素养。教研室全体教师不仅具备扎实的信息检索理论知识,而且还拥有丰富的信息咨询、查新查引、文献传递等实际工作经验,更重要的是能了解读者的需求和心理,将理论与实践完美地结合在教学过程中,所授课程得到了学生的一致好评。

多年的教学研究和实践凝结成了这部教材。本教材编排体例科学,内容结构完整、系统,语言表达准确、精炼,突出知识要求和检

索技巧与方法的讲解,对各检索系统的特点进行分析,是一本面向高校学生的综合性、实用性教材。文中还采用了大量的实例图表,帮助读者更直观地理解知识点和掌握各类检索系统的使用方法,力争做到内容系统完整、阐述科学简明、技术新颖实用。

本教材由王宇主编,主要负责结构体例设计、审校与统稿工作,并承担部分章节的撰写工作。吴瑾、刘允斌、官平、张英敏、洪跃、王贵海、丁学淑、芦金梅、董绘丽、胡永强、王磊、刘偲偲等参与编写。

在本教材内容的组织、整理及写作的过程中,参阅了大量的相关文献,借鉴了同行学者的有关论著和研究成果,对此,我们向有关人士表示衷心的感谢!我们诚挚地希望本教材能够对读者尤其是高校学生读者的信息检索和利用能力的提高有所帮助。

由于信息检索技术的不断发展及检索手段的不断更新,加之编者时间和知识水平有限,书中难免有不当之处,敬请专家、同仁及读者批评指正。

编 者

2013年1月

目 录

基础篇

1 信息检索绪论	003
1.1 信息、知识、文献及情报的基本概念	003
1.2 信息的性质与分类	008
1.3 信息检索的目的、意义与作用	011
2 信息检索的基本知识	016
2.1 信息检索的概念与类型	016
2.2 信息检索语言	019
2.3 信息检索的方法与步骤	029
2.4 检索效果评价	040

应用篇

3 图书信息检索	045
3.1 书目信息检索	045
3.2 中文电子图书检索	051
3.3 外文电子图书检索	061
4 期刊信息检索	063
4.1 中文期刊检索	063
4.2 外文期刊检索(文摘型)	093

4.3 外文期刊检索(全文型).....	108
5 网络信息检索	122
5.1 网络基础知识.....	122
5.2 网络信息资源.....	128
5.3 网络信息检索工具.....	130
5.4 开放存取(Open Access)	154
6 特种文献信息检索	169
6.1 学位论文信息检索.....	169
6.2 会议文献信息检索.....	175
6.3 专利文献信息检索.....	179
6.4 标准文献信息检索.....	189
6.5 科技报告信息检索.....	198

综合篇

7 文献传递与数字参考咨询服务	205
7.1 文献传递	205
7.2 数字参考咨询服务	217
8 文献信息管理与利用	221
8.1 文献管理软件	221
8.2 科研课题与科技查新	237
8.3 学术论文写作与投稿指南	243
8.4 网络学术论坛	251
参考文献	256

基 础 篇

1.1 信息、知识、文献及情报的基本概念

1.1.1 信息

从古至今,人类的知觉让生活一直与信息密切相关。人类通过了解自然的信息来认识自然,通过社会的信息来知悉人类社会的过去,利用信息进行发明创造。人类一直在不知不觉中创造、利用着信息。今天,我们已经进入了信息社会,信息包围着我们的生活,可以毫不夸张地说,我们今天是生活在信息的海洋中。“信息”、“信息社会”、“信息时代”等名词成了社会使用频率非常高的名词。的确,信息作为一种现象,它广泛地存在于自然界和人类社会中,信息和物质、能量是构成客观世界的三个基本要素。信息无处不在,无人不用,今天我们的生活已经离不开信息。然而,究竟什么是信息?自从1948年美国应用数学家申农创立信息论以来,信息概念一直是理论界关注的焦点,也是信息时代争议最多的难点和最具基础性、挑战性的核心概念之一。

关于信息的定义可谓众说纷纭,目前有很多种,人们常把信息理解为具有新内容、新知识的消息。在《辞海》中,信息就被解释为“对消息接受者来说预先不知道的报道”,这种说法有一定的道理,但它并没有揭示出信息的本质和最实质性的内容。美国数学家维纳在1948年发表的《控制论——动物与机器的通讯与控制问题》中指出:“信息既不是物质,又不是能量,信息就是信息。”他在另一部著作《控制论和社会》中,对信息下的定义是:“信息是我们适应外部世界并且使这种适应为外部世界所感知的过程中,同外部世界进行交换的内容的总称。”而信息论的首创者申农于1948年发表的《通讯的数学理论》一文中指出,一个系统所接收的“信息是能够用来消除不确定性的东西”。美国《未来学家》杂志所刊登过的专家们撰写的文章中指出,信息是一切容易获得的和不易获得的、有时可供人们参考的事实和思想的总和。

这些观点虽然都对信息的本质有了一定程度的揭示,但他们都没有给信息一个明确的定义。我国信息理论界学者用辩证唯物主义的观点,对信息作出了种种解说,其中较有影响的有:

“信息是事物存在方式或运动状态,以及这种方式(状态)的直接或间接的表述。”(钟义信:《信息科学》)

“信息是标志物质间接存在性的哲学范畴,它是物质存在方式和状态的自身显示。”(邬焜:《哲学信息论要略》)

“信息是物质的普遍属性,它表述其所属的物质系统,在同任何其他物质系统全面相互作用的过程中,以质、能波动的形式所呈现的结构、状态和历史。”(黎明:《信息时代的哲学思考》)

从以上信息定义可以看出,作为信息的概念起码应包含这样三个基本的内容层次:信息是物质的普遍属性,是物质的存在方式;信息是物质的存在方式、状态的表述,或可以说信息是物质存在方式和状态的自身显示;信息产生的动力存在于物质自身相互作用与联系之中。

据此,我们把信息定义为:信息是物质的一种存在方式或状态,以及这种方式、状态的直接或间接的显示或表征,信息是事物本质、特征和运动规律的反映。

1.1.2 知识

在生活中,我们经常会听到很多关于知识的名言,例如,培根:“知识就是力量”;高尔基:“爱书吧,它是知识的源泉”;爱因斯坦:“学习知识要善于思考,思考,再思考”;邓小平:“靠空讲不能实现现代化,必须有知识,有人才。没有知识,没有人才,怎么上得去。”这么多关于知识的名言,但我们是否真正理解什么是“知识”呢?如同信息一样,古往今来,人们对知识的理解也是仁者见仁,智者见智的。

知识是一个内涵丰富的概念,有关知识的定义有很多,如:

《辞海》1989年版解释为:人类认识的成果或结晶。

《现代汉语词典》1991年版对知识的定义是:人们在改造世界的实践中所获得的认识和经验的总和。

古希腊:所有知识都是来自个人的经验,而每个人对事物的感受不同,因此知识仅仅是相对于个人而言的,没有绝对意义的知识。

柏拉图:知识是经过证实的正确的认识,只有在正确方法论指导下获得的认识才是知识,否则只能叫想法或信仰。

德鲁克:知识是一种能够改变某些人或某些事物的信息,经过人的思维整理过的信

>>>>>>
息、数据、形象、意象、价值标准以及社会的其他符号化产物等。

人们通过来自自然界和人类社会的不同信息以区别各种事物,从而认识世界,并最终改造世界,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、汇集成知识。人类社会的进步,正是知识不断积累、不断更新的过程。进一步讲,知识实质上就是对世界发出的信息加以接收、存贮、加工、归纳中得到的产物。随着这种反映的逐步深入,人们的认识不断深化,从而导致认识的飞跃——从感性认识到理性认识,这种认识飞跃的结果就是知识。

综上所述,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,知识是信息的一部分,是被人们理解和认识并经头脑重新组织和系列化了的那部分信息。

在知识的产生和发展过程中,逐渐表现出了知识的一些属性,主要有:

信息性。信息是生产知识的原料,知识是通过人类认识、理解并经思维重新整合后的系统化信息,知识是信息中的一部分。

意识性。知识是一种观念形态的东西,我们在头脑中产生它,随着认识的深化,我们就在生活中利用它。

规律性。人们在实践中对事物的认识,是一个无限的过程,人们在这种无限过程中所获得的知识从一定的层面上揭示了事物及其运动过程的规律性。

实践性。实践是产生知识的基础,也是检验知识的标准,知识又对实践具有重大的指导作用。

继承性。对于新知识的产生,既是原有知识的继承、利用、深化与发展,又是下次知识更新的基础和前提。

渗透性。随着人类认识世界的不断深化,各种门类的知识可以相互渗透,构成知识的网状结构。

1.1.3 文献

文献是信息存在的两种方式之一,并且是一种非常重要的存在形式。人们为了将各种实用的知识信息储存、传递下去,就需要载体,这个载体就是文献。结绳记事就是最原始的文献萌芽。真正的文献发展和丰富则是在文字产生以后,而当今文献在内容和形式上与古代的文献都有很大的不同。首先是文献的载体不同:从甲骨、竹简、羊皮、纸莎草、石(泥)板、丝帛、纸张发展到用感光介质和磁性介质,直至现在的网络;其次是文献的表达形式和手段也不相同:人们认识到文字不再是表达思想的唯一手段,图形、声频、视频都是表达思想、传递感情的重要手段。国际标准化组织《文献情报术语国际标准》对文献的解释是:“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或数据的载体。”

文献的构成具有一定的要素,现代文献由四个基本要素构成,即文献信息、物质载体、符号系统和记录方式。文献信息是指文献所记载的内容;物质载体是符号赖以存在的“寄主”,即文献的内容记录在这些形式的物质之上,有纸质文献、缩微文献、音像文献、磁盘和光盘文献等;符号系统包括图画、文字、图表、编码、声像和电磁等信息,它是信息的撰写者;而记录方式则是指文献符号以何种方式被记录下来,是指将表达信息的符号系统通过特定的人工记录方式使其附着于一定的文献载体上。记录的方式有很多,有写画、雕刻、印刷、摄影、扫描、拷贝等。人类自有史以来,从对世界的认识中形成了极为丰富的知识,但被人们通过各种载体用文献记录下来的只是其中的一部分。

从上述定义不难看出,文献是指记录有知识或情报的一切载体,知识或情报要依附在一定的物质载体上才能具体体现出来,并得以长期保存和广泛传播。

文献的类型有多种多样,不同的标准划分的结果不同,其中根据文献的载体类型可以将文献划分为:① 印刷型:是文献载体的最基本方式,包括铅印、油印、胶印、石印等各种资料。优点是可直接、方便地阅读。② 缩微型:是以感光材料为载体的文献,又可分为缩微胶卷和缩微平片,优点是体积小、便于保存、转移和传递,但阅读时须用阅读器。③ 机读型:是一种最新形式的载体。它主要通过编码和程序设计,把文献变成符号和机器语言,输入计算机,存储在磁带或磁盘上,阅读时,由计算机输出。它能存储大量信息,可按任何形式组织这些情报,并能以极快的速度从中取出所需的情报。近年来出现的电子图书即属于这种类型。④ 声像型:又称直感型或视听型,是以声音和图像形式记录在载体上的文献,如唱片、录音带、录像带、科技电影、幻灯片等。

文献在科学和社会发展中所起的作用,主要表现在:① 它是科学研究和技术研究结果的最终表现形式。② 它是在空间、时间上传播情报的最佳手段。③ 它是确认研究人员对某一发现或发明的优先权的基本手段。④ 它是衡量研究人员创造性劳动效率的重要指标。⑤ 它是研究人员自我表现和确认自己在科学中的地位的手段,因而是促进研究人员进行研究活动的重要激励因素。⑥ 它是人类知识宝库的组成部分,是人类共同财富。

可以看出,文献在人类社会发展的进程中,发挥着举足轻重的作用,它是古代文明的传承者,它是改革创新推动者。

1.1.1.4 情报

在古代,情报首先产生于军事领域,《辞源》指出:“定敌情如何,而报于上官者”为情

报;《辞海》中则说“战时关于敌情之报告,曰情报”,这是中国最早期的情报定义。它反映了情报作为消息传递的功能及构成情报的两个基本要素——“情”与“报”,强调情况、消息的传递报道作用。到了近代,随着科学技术的迅速发展,创造与传播知识的工作有了新的发展,专职情报机构的主要工作是使知识有序化,以解决问题。于是,情报的概念也有了新发展,认为情报是“作为存储、传递和转换的对象的知识”。为了满足用户的多种需要,有人从特定概念出发提出:“情报是在特定时间、特定状态下,对特定的人提供的有用知识。”情报资料激增影响对决策人员的有效服务,为了解决这个问题,情报工作开始由一般文献工作阶段进入了侧重与经济、社会发展相结合的情报分析研究阶段,情报的定义增添了新的内容:“情报是判断、意志、决心、行动所需要的能指引方向的知识 and 智慧”,是“解决问题所需要的知识”,“情报是激活了的知识”。情报来源于人类社会实践,是物质世界与精神世界共同作用的产物。人类正是在不断认识、改造自然与社会的过程中,在物质生产与科学实验的实践中,源源不断地创造、交流与利用各种各样的情报。在日常生活中,人们经常在不同的领域里,自觉或不自觉地在传递情报、接收情报与利用情报。因此,情报又是一种普遍存在着的社会现象。综上所述,我们可以把情报简单地解释为:传递中的有用的新的知识。

情报具有三个基本属性:知识性、传递性与效用性。知识是人的主观世界对客观世界的概括和反映。随着人类社会的发展,每时每刻都有新知识产生,人们通过读书、看报、听广播、看电视、参加会议、参观访问等活动,都可以吸收到有用的知识。这些经过传递的有用知识,按广义的说法,就是人们所需要的情报。因此,情报的本质是知识,没有一定的知识内容,就不能成为情报。知识性是情报最主要的属性。但知识成为情报,还必须经过传递,知识若不进行传递交流,供人们利用,就不能构成情报。所以,情报的传递性是情报的第二基本属性。第三是情报的效用性,人们创造情报、交流传递情报的目的在于充分利用,不断提高效用性。情报的效用性表现为启迪思想、开阔眼界、增进知识、改变人们的知识结构、提高人们的认识能力、帮助人们去认识和改造世界。情报为用户服务,用户需要情报,效用性是衡量情报服务工作好坏的重要标志。此外,情报还具有社会性、积累性、与载体的不可分割性以及老化等特性。情报属性是情报理论研究的重要课题之一,其研究成果丰富了情报学的内容。

对于情报的分类,依据的原则和采用的方法不同,得到的结果也不相同,如按情报的应用范围可分为科学情报、经济情报、军事情报、政治情报等;按情报的内容及其所起的作用,又可分为战略性情报和战术性情报两大类;按自身的性质和所处的阶段,又可分为固有情报、接收情报和有用情报。

1.2 信息的性质与分类

1.2.1 信息的性质

信息具有多方面的性质,从不同角度去认识和理解,就会表现出不同的特性,主要有以下几个方面:

1. 信息的普遍性和客观性

信息是事物变化和状态的客观反映,它和所反映的客观事物一样,其状态、特征和变化具有不以人的意志为转移的客观性。

2. 信息具有可传递性和扩散性

现代信息技术的发展,特别是全球“信息高速公路”的兴建,将会使信息传递和扩散的速度越来越快、方式更为便捷,让人们更加方便地获取所需的信息。

3. 信息具有价值性

信息是一种资源,是和物质、能源并举的三大资源之一,这是人们的基本共识。信息的价值性表现为信息效用,能够满足人们的信息需求,造福于人类社会。所以,我们应该提倡人们在获取信息与信息需求之间建立沟通的渠道,以最大限度地实现信息的价值。

4. 信息具有时效性

信息有自己特定的生命周期。时效性与价值性之间,具有双向建构关系。信息的使用价值往往与其所提供的时间成反比,时间的延误,会使信息的使用价值衰减甚至完全消失。现代科学知识老化周期的加快,社会竞争意识的日渐强化,人们越来越感觉到:时间就是金钱,信息更是金钱。

5. 信息的可共享性

信息的可共享性是相对于物质的独占消费特征而言,这也是信息交流与物质交流的本质区别。但信息的可共享性不等于信息利用的无偿性。世界各国在大力倡导信息资源共享的同时,更注重知识产权的保护。

6. 信息与载体的不可分割性

人类社会的各种信息活动所产生的信息,必须借助于文字、图像等工具和某种物质形态的载体,才能够得以表现和存在,并被人类自身所感知和认识,体现出信息的价值,否则就不成为信息,至少是无用的信息。

7. 信息具有可存贮性或再现性

这是由信息的客观性和信息与载体的不可分割性派生出来的基本特性。正是借助这一基本特性,使人类文化的继承和延续能够得以实现。

8. 信息具有累积性、延续性和继承性

这也是由可存贮性和再现性派生出来的信息特征。人类知识宝库的不断丰富和扩充过程,就是信息经过系统化、抽象化和规律化而形成知识的长期积累过程。直接探索、认识和获取一个未知事物的信息是极其困难的,而接受、理解和继承一个信息则简单、容易得多。

1.2.2 信息的分类

信息的分类和信息概念本身一样,没有统一的标准。人们试图从不同的标准和分类依据对信息的种类加以归纳,以帮助我们认识信息的内涵。我们认为通常采用的分类方法有以下四种:

1. 按照信息的内容划分

自然信息。又称非人类信息,是指人类以外的宇宙天体、自然界万事万物所产生的和彼此交流的信息。主要包括与人类社会无关但对人类社会活动有用的生物信息和非生物信息,以及人类社会暂时还无法利用的无用信息。

文化信息。又称人类信息,是指人类从事各种社会活动所产生的彼此交流的信息。一切经过人类利用语言、文字、符号、图像等方式加工过的自然信息均称为文化信息。知识、情报、指令、告示、法律、各种艺术形式、神话、传说……这些全都属于文化信息。文化信息内容是社会生活的反映,具有鲜明的目的性和有用性,远比自然信息丰富多彩。

2. 按信息的性质划分

可分为语义信息和非语义信息。语义信息是指用语言、文字或语义性质的符号所描述的信息;非语义信息是指用非语义性质的声音、颜色或实物荷载的信息。语义信息为我们人类所感知,非语义信息中相当一部分也是为我们人类可感知的,并对其含义有所认识和理解的信息。

3. 按照信息产生的先后和加工程度划分

零次信息。是指未经任何修饰加工的源信息,源信息产生于人类直接的政治、经济、文化等活动,是一种零星的、分散的、无规则的信息。

一次信息。又称原始文献信息。是人们研究或创造性活动成果的直接记录,如科技报告、会议论文、学位论文、期刊论文、图书等。所记录的信息一般比较具体和详尽,有很高的直接参考和借鉴使用价值,反映人类科学文化发展的直接成就,是人类文明的财富和

象征,是信息检索的主要对象和最终目标。

二次信息。是对一次信息进行有序化整理后的信息,它具有便于存贮、检索、传递和使用等特征,能为获取一次信息提供具有针对性的线索和途径。

三次信息。三次信息是指根据二次信息提供的线索,查找和使用一次信息以及其他材料,进行浓缩、整合后产生的信息,如专题评述、研究报告、动态综述、百科全书、年鉴、数据、手册等。它集中反映了科学研究与发展的最新前沿成就和状况,是任何具有敏锐信息意识的人士都十分关心和迫切需要获取的对象。

4. 按照信息的载体形式划分

感官载体信息。又称主观知识信息,它以人的大脑为载体,是客观事物对脑神经的刺激所产生的反映。

语言载体信息。是指借助语言进行交流和传播的信息,又称口头信息。

文字载体信息。又称文献载体信息,是信息的主要存在形式,也是传统的信息记载方式。

缩微载体信息。其载体是胶片、照片,由于它一般是以文献载体为蓝本缩微而成的,因此,有时把它称为文献载体的特殊形式。

其他载体信息。如电磁波载体信息、光波载体信息、声像载体信息、电子计算机载体信息等,代表着未来信息载体的发展趋势,其优势表现为体积小、贮存容量大、保存时间长、使用方便。

5. 按照信息的出版形式划分

这是一种常见的分类形式,可分为图书、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。

信息分类的方法还有很多种,采用不同的分类标准,就可以从不同的角度或某一方面对信息进行分类,例如,依据信息的活动状态,可以将信息分为静态信息和动态信息;按信息资源组成与内在关系划分可分为元信息资源、本信息资源、表信息资源;按信息资源的有形与无形可分为有形信息资源与无形信息资源;按其地域(所处空间范围)可分为国际(世界)信息资源、国家信息资源、地区(部门)信息资源、单位信息资源;按信息传递的范围可分为公开信息资源、半公开信息资源、非公开信息资源等。

1.2.3 信息、知识、文献及情报之间的关系

1. 信息、知识、文献及情报之间的从属关系

信息、知识、文献、情报之间在内容和范围上存在一定的包含关系(图 1-1)。信息是物质的一种存在形式,和其他客观事物一样,无处不在,无时不有。知识、文献和情报都是