



现代信息 检索与利用

主编 陶颖

黑龙江教育出版社

现代信息检索与利用

主 编 陶 颖

副主编 孙 明 周 莉

彭 焱 吕光远

黑龙江教育出版社

图书在版编目(CIP)数据

现代信息检索与利用/陶颖主编. —哈尔滨:黑龙江教育出版社,2006.9

ISBN 7-5316-4645-5

I. 现... II. 陶... III. 情报检索 IV. G252.7

中国版本图书馆 CIP 数据核字(2006)第 116166 号

现代信息检索与利用

主 编 陶 颖

副主编 孙 明 周 莉

彭 焱 吕光远

责任编辑/杨雪松

封面设计/神龙公司设计中心

责任校对/李宇

出版发行/黑龙江教育出版社(哈尔滨市南岗区花园街 158 号)

印 刷/哈尔滨市工大节能印刷厂

开 本/787×1092 毫米 1/32

字 数/178 千

印 张/8

版 次/2006 年 9 月第 1 版

印 次/2006 年 9 月第 1 次印刷

书 号/ISBN 7-5316-4645-5/G·3542

定 价/32.00 元



作者简介

陶颖

1959年生人

现任哈尔滨商业大学图书馆馆长

毕业于哈尔滨师范大学

副研究馆员

在多年从事图书馆管理工作过程中

潜心研究

在图书馆管理

信息检索等方面颇有建树

在《中国图书馆学报》

《情报资料》

《管理世界》等刊物发表专业

论文10余篇。

目 录

第一章 绪论	(1)
1.1 信息	(1)
1.2 信息素养	(11)
第二章 信息检索基础知识	(17)
2.1 信息源	(17)
2.2 信息检索的概念和基本原理	(25)
2.3 信息检索途径	(28)
2.4 信息检索语言	(30)
2.5 信息检索工具	(35)
2.6 信息检索方法	(39)
2.7 信息检索技术	(40)
2.8 信息检索的基本程序	(44)
2.9 信息检索效果评价	(47)
2.10 文献信息分布规律	(51)
第三章 参考工具书及其网络资源	(56)
3.1 工具书概述	(56)
3.2 参考工具书及其网络资源	(61)
第四章 特种文献信息检索	(83)
4.1 专利信息检索	(83)
4.2 学位论文检索	(100)
4.3 标准信息检索	(110)

第五章 网络信息资源的检索与利用	(125)
5.1 Internet 概述	(125)
5.2 国内主要网络简介	(130)
5.3 网络信息资源	(131)
5.4 网络信息资源检索	(136)
第六章 网络信息检索工具	(143)
6.1 网络信息检索工具概述	(143)
6.2 网络信息检索工具的类型	(145)
6.3 搜索引擎	(150)
第七章 网络数据库检索	(173)
7.1 网络数据库概况	(173)
7.2 中文网络数据库	(176)
7.3 国外网络数据库	(205)
第八章 电子图书、电子期刊、电子报纸的检索	(212)
8.1 电子图书 E-book	(212)
8.2 电子期刊	(220)
8.3 电子报纸	(222)
附录 1 中国图书馆分类法简表	(226)
附录 2 Internet 上部分重要学术信息网址参考	(245)
参考文献	(252)

第一章 绪 论

信息与能源、材料并称为现代社会的三大支柱。20 世纪以来,信息高速增长,在人类社会中日益发挥重要作用。可以说,信息从来没有像今天这样成为人们争夺的最重要的经济资源、战略资源和社会资源,它已经成为维系人类社会生存和发展的基础,成为衡量一个国家综合实力和竞争力的重要标志。

作为 21 世纪的现代人,必须具备信息素养,必须具备信息检索知识,这是现代人生存和发展的重要因素。信息是智慧的源泉,信息是财富,而信息检索知识以及信息素养则是打开信息这座宝库的钥匙。

本章将介绍信息的基本知识和信息检索的重要性。

1.1 信息

1.1.1 信息的定义

人类生活在信息社会,每时每刻都在自觉或不自觉地与信息打交道。那么究竟什么是信息呢?这是一个难以回答和有争议的问题。一方面是因为信息术语作为一个抽象的概念,不同的学者对它有不同的认识;另一方面是因为不同学科领域对信息术语有不同的理解,赋予它不同的意义。

信息作为一个科学术语最早出现于通信领域,20 世纪中叶被引入哲学、信息学、系统论、控制论、情报学、经济学、管理学、计算机等领域。目前,关于信息的定义众说纷纭,莫衷一是。

●美国《韦氏词典》:信息是用来通信的事实,在观察中得到的数据、新闻和知识。

●英国《牛津词典》:信息就是谈论的事情、新闻和知识。

●日本《广辞苑》:信息是所观察的事实和知识。

●我国《辞海》:音讯;消息。早在一千多年前,唐朝诗人李中在《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”的诗句。通信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。通常必须通过处理和分析来提取。信息和物质、能量被称为构成系统的三大要素。信息的量值与信息的随机性有关,如在接收端无法预估消息或信号中所蕴涵的内容或意义,即预估的可能性越小,信息量就越大。

●美国《未来学家》杂志:它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。

●我国图书馆学界认为:“信息不是事物本身,而是由事物发出的消息、情报、指令、数据、信号等当中所包含的内容。”

上述定义把信息视为具有普遍性的事物运动的状态和方式,但是没有提及信息的本质。

●控制论的奠基者维纳认为:信息这个名称的内容,就是我们对外界进行调节并使我们的调节为外界所了解时与外界交换来的东西。接收信息和使用信息的过程就是我们对外界环境中的各种偶然性进行调节并在该环境中有效地生活的过程。

他把信息定义为:信息是人们在适应外部世界并使这种适应反作用于外部世界的过程中,同外部世界进行交换的内容的总称。信息就是信息,不是物质也不是能量。这个定义说明信息是区别物质、能量的另一个构成世界的重要因素。

●信息论创始人申农把信息系统定义为:由信源、信道和信宿组成的系统。

信息具有使不确定性减少的能力。信息量就是不确定性减少的程度。申农认为,进行通信无非是通信的信源一方有某种形式的消息要告诉对方,并估计到对方对此存在着不确定性。申农把信宿获得信息看作是可使其不确定性减少和消除。

●我国学者钟义信:信息是事物存在的方式或通信的状态,以及这种方式、状态的直接或间接的表述。

●我国学者严怡民:信息是生物以及具有自动控制系统的机器,通过感觉器官和相应的设备与外界进行交换的一切内容。

从以上解释可以看出,信息的概念是有层次的。主要有两大层次:一是抽象的理论意义上的信息概念,二是具体的实用意义上的信息概念。从本体论层次上看,信息是一种客观存在的现象,是事物的运动状态及其变化方式;从认识论层次上看,信息是主体所感知或所表述的事物运动状态及其变化方式,是反映出来的客观事物的属性。而具体的实用信息概念是从不同的利用角度以及不同的学科角度对信息的界定,有其特定的内涵和外延。

1.1.2 信息的属性

越是基本的概念,越难定义。为了更好地理解信息的概念,我们可以从信息的属性入手进行研究。下面主要探讨一下信息的存在属性、使用属性、价值属性。

1. 信息的存在属性

(1) 信息的普遍性和客观性

世界是物质的,物质是运动的,而信息是事物的运动状态及其变化方式,物质和运动的普遍性和客观性决定了信息的普遍性和客观性。

(2) 信息的依附性和多态性

信息不是事物本身,任何信息都依附于特定的载体而存在。载体的多样性决定了信息的多态性;人类对信息的描述语言、表述

方式也决定了信息的多态性。

(3) 信息的可包装性

信息可以进行分类、加工、整理,可以压缩、伪造,这就是信息的可包装性。

(4) 信息的可扩散性或可传递性

信息的可扩散性也叫信息的可传递性。表现在时间和空间两个方面。在空间方面,信息从浓度高的地方向浓度低的地方流动,距离信息源越近越容易获得信息。信息在时间传递方面表现为过去的科学规范变为今天的科学常识,随着人们认知水平的不断提高,信息反映出的内容也日益深刻,推动社会的发展。

2. 信息的价值属性

信息就是传递中的知识差。信息的价值在于它的知识性,在于它可以消除认识的不确定性。不确定性实际上是人们对于知识的无知程度,或者说人们对环境状态的无知程度。某种原因造成的不确定性会给人们决策活动带来风险。面对不确定性的风险,很重要的一点就是要充分地获取信息,以减少不确定性本身。

3. 信息的使用属性

(1) 信息的无损使用和增殖传输性

与物质和能量不同,信息是不会随着使用而消耗的,不会在传输中损耗和消亡。当多个信息汇聚起来,有可能产生新的信息和知识。

(2) 信息的聚变效应和可繁衍性

关于信息的作用有一条特殊的规律:全部大于部分之和。一方面人们进行决策时,收集到的信息越多,汇集起来的作用越大,越容易进行正确决策;另一方面,人们通过互相交流,认识会得到深化,知识面会得到发展。这些都是因为信息是可以聚变繁衍的。信息收集得越多,所繁衍的信息也就越多,信息内容的深度越大,

越容易剔除无用或错误信息,提高质量。

(3)信息的共享性

信息可以供不同的系统分享,而不计分享者多少。这是信息有别于物质和能源的重要属性之一。

(4)信息的适合性

也称信息的可连接性。它主要表现为信息强度、形式、内容的适合性,无论对自然系统还是对人,都一样。例如收音机只能接收一定强度、一定频率范围内的电磁波。

信息对人的适合性表现在很多方面。人在接受信息的过程中,受到很多因素的影响。接受信息时,人的感官和神经处于紧张状态,所接受的信息强度越大,感官和神经越容易疲劳。当人处于疲劳状态时,接受信息的能力就会削弱,所以人们应该学会张弛有度。不同的信息形态适合人的不同感官;不同的感官有不同的接收效率。据统计,在人类通过感觉器官所接收的全部信息中,视觉约占65%,听觉约占20%,触觉约占10%,味觉约占2%,大部分信息来源于视觉和听觉。当人用多个感觉器官接收同一内容的信息时,要比用单一的感觉器官接收效率高得多。因为种种原因信息具有一定的冗余以及多态性,一个感官休息时,另一些感官还在工作,交替使用,不易疲劳。

抽象信息比形象信息需要经过大脑较多的处理,因而接收的效率低且更容易引起疲劳。信息形态也影响信息的适合性。

个人兴趣可以引起大脑不同程度的兴奋。需要、兴趣、知识水准可以决定人的感官与信息的匹配程度。当信息与己无关时,会使匹配程度降低,甚至关闭感官。如对一名没有接触过数学的小学生讲大学的微积分,因为存在知识的不连接性,他是无法接受更高级的知识的。

(5)信息的不完全性

对于信息不完全性的研究是从市场信息的特点开始的。早期的经济学家把市场假定为完全信息市场,即以市场的参加者对某种经济环境具有全部知识为前提。新古典经济论认为,在市场上,消费者在每一个时点上都了解市场各种商品的全部可能价格,也清楚自己的偏好、存货;厂商也知道生产要素、价格和投入产出的全部知识以及各种形式的组合配置。

如果市场是完全信息市场,就没有信息差,信息也就没有存在、传播的必要了。

(6) 信息的不对称性

信息的不对称性是博弈论研究的内容。市场信息的不对称是相对于市场信息的对称性而言的。所谓对称市场信息,是指在市场上,参加交易的双方——顾客和厂商要么双方什么都不了解,要么双方都具有完全信息,要么双方具有的不完全信息是完全相同的。但是在现实社会中,这是不可能的。个人的信息能力也能造成信息不对称。

1.1.3 信息的分类

信息广泛存在于自然界和人类社会,具有广泛的内涵。从不同角度,按不同的标准,信息可以划分成很多种类型。常见的信息分类有:

从信息学科内容 and 应用角度可划分为政治信息、科技信息、文化信息、经济信息、军事信息、教育信息等。

按信息的感知方式可划分为直接信息和间接信息。

按信息所依附的载体可划分为文献信息、口头信息、电子信息、生物信息等。

按信息的记录符号可划分为语音信息、图像信息、文字信息、数字信息等。

按信息的运动状态可划分为连续信息和离散信息。

按信息的加工层次可划分为初始信息和再生信息。

按信息的可开发程度可划分为潜在信息和现实信息。

1.1.4 信息的作用

1. 信息是社会发展的资源

现代社会已经把信息与材料、能源视为支持社会发展的三大支柱,这充分反映出信息在现代社会中的重要作用。随着现代技术在全球的应用和发展,信息已成为企业创造财富最主要的资本,成为生产增长的内生变量。信息产业已经成为国民经济的先导产业。世界上许多资源贫瘠的国家如日本依靠“信息的力量”成功地跻身世界经济强国之列。日本能够在二战后崛起,靠的就是从国外引进的先进技术资源、信息通畅,所以有人说日本是靠信息起家的国家。

2. 信息是社会交流的工具

从信息科学的观点来看,信息是社会的联结方式。人类的活动并不是孤立的个体活动,它通常表现为以个人活动为基础的社会性。人类活动正是因生存的需要形成了社会,而人与人之间的信息交流是人类活动社会性赖以形成和发展的根本保证。控制论的创始人维纳称信息是人类社会的黏合剂,人类在漫长的进化过程中,在每一阶段的生存、延续和繁衍都伴随着信息交流;没有交流,人类的进化和生存就无从谈起。人类正是由于自身所具有的信息能力,才能不断地从自然界接收信息,与同类进行信息交流,不断地维系生命、繁衍生命、扩大自己的生存空间。

3. 信息是人类认识世界和改造世界的媒介

人类的思维能力使人能够认识世界和改造世界。思维能力是人的认识能力的核心,而思维功能的实现则是基于信息进行的。如果没有信息、没有知识,人的思维能力也就无从谈起。信息是思维的原材料,也是思维产生的结果。

信息是维系人类与自然界关系的媒介,是人类认识世界的必要条件。人类不仅能认识世界,而且能改造世界,在提高生产力的同时,最大限度地创造财富,积累和发展物质文明和精神文明。人类改造世界的能力与人类的信息能力密切相关。只有具备较强的信息能力,充分利用信息,才能更好地认识和改造世界。

4. 信息是决策的依据

决策是人类社会实践活动的重要环节,对个人、组织乃至国家都有十分重要的意义。信息是决策的依据,是决策的关键因素,是决策者与决策对象之间的桥梁,在决策活动中起着重要的作用。决策过程中的每个环节都离不开信息。决策者要从信息中发现问题,更要利用信息解决问题。只有掌握了大量信息并善于对信息进行分析研究,才能做出科学合理的决策。

1.1.5 信息的相关概念——知识和情报

知识、情报这两个概念与信息密切相关。

1. 知识

知识是人类对客观事物存在和运动规律的认识。

1989年版《辞海》解释为:人类认识的成果或结晶。包括经验知识和理论知识。经验知识是知识的高级形态。人类的知识是在后天社会实践中形成的,是在社会实践中积累起来的经验,是对现实的反映。辩证唯物主义把社会实践作为一切知识的基础和检验知识的标准。知识借助于一定的语言形式,物化成为某种劳动产品的形式,交流和传递给下一代,成为人类共同的精神财富。知识随社会实践、科学技术的发展而发展。知识一般可以分成三大类:自然科学知识、社会科学知识和思维科学知识。哲学知识则是关于自然、社会和思维知识的概括和总结。

也有人认为,知识是人类在改造客观世界实践中所获得的认识和经验的总和。人们认识客观事物的过程就是人脑对外界事物传

来的信息进行加工的过程。而认识飞跃的结果即为知识。也就是说信息被有选择地吸收到人们的思维系统,经过人们大脑的储存、甄别、加工、处理、转换等形式而形成知识。知识是信息的一部分,信息的概念要大于知识。人们通过信息感知世界、认识世界和改造世界,并且将获得的信息转变成知识作为认识和改造世界的武器。我们可以说知识是一种信息,而不能说信息就是知识。知识与信息可以相互转化。使用者把信息转化为知识,再把知识转化为智慧,是一种动态过程,是一种开拓过程。反过来,智慧又会转化为新知识,新知识又会转化为新信息,人们通过一定的社会传递过程或媒体把信息传递给使用者。

知识可分为理性知识和感性知识。理性知识是对客观事物的本质和规律性的认识,是经过思维、逻辑加工的信息。感性知识是对客观事物的描述和对现象、事实的感知,是未完全经过逻辑加工的信息内容。

经济合作与发展组织(OECD)出版的《以知识为基础的经济》报告将知识分为四种类型:

(1)知道是什么的知识(Know - what)。指关于事实与现象的知识,是一切知识的基础。包括自然科学知识和社会科学知识,它可以通过学习获得。

(2)知道为什么的知识(Know - why)。指那些自然、人的思维、社会运动的法则和规律的科学知识,包括自然原理和规律方面的科学理论。

(3)知道如何做的知识(Know - how)。指技艺和能力知识,包括技术诀窍或专有技能。

(4)知道是谁的知识(Know who)。指知道谁知道什么、谁知道如何做的知识,包括与有关专家形成的特殊社会关系,以便有可能获得并有效利用这些专家的知识,侧重对创造思想、方法、手段、

过程以及特点等的了解。

2. 情报

情报与信息关系密切,它们的英文单词都是“Information”,但我国学者普遍认为二者是有区别的。

情报是一个发展中的概念。在不同时期,其社会功能和概念都有所不同。古代的情报用于军事领域,近代之后人们对情报加深了认识。现代有人认为情报是人和人之间传播着的一切符号系列化的知识,是作为存储、传递和转换的对象的知识。我国情报界较认同的定义是:情报是作为交流对象的有用知识或消息,也可以说是用于特定需要的知识或消息。与信息概念一样,人们对情报的认识也是多种多样的。无论怎样,情报将随着社会功能的扩展而增添新的含义,新的情报概念是在旧有概念的基础上发展而成的。

从情报的定义来看,情报也是人类社会的一种普遍现象,它存在于人类社会的各项活动之中。从这些意义上讲,情报与信息具有一致性。

二者的不同点表现在:情报作为科学交流的对象是指一切传递中的能够改变认知主体原有知识结构的知识,它才是情报概念的主体构成部分。我国著名科学家钱学森指出:“情报就是为了解决一个特定的问题所需要的知识,要注意它的及时性和针对性这个要求”;“情报就是激化了、活化了的的知识”。英国著名学者、情报学家布鲁克斯准确也更科学地描述了情报与知识之间的相互作用关系。他曾提出这样一个基本方程式:

$$K[S] + \Delta I = K[S + \Delta S]$$

其中, $K[S]$ 表示原有的知识结构; ΔI 为新吸收的情报量; $K[S + \Delta S]$ 为新形式的知识结构。

该公式充分表明,没有情报的传递和交流,就不会有知识的利

用和吸收;情报是知识的一部分,但静止的知识并不是情报,情报的价值正在于它是交流着的知识。由此可见,情报就是传递着的知识,而且这种传递着的知识是可以改变认知主体的知识,从而解决其特定问题。

另外,一种经过加工、分析之后产生的“情报”,其含义既不同于信息,也不同于一般意义上的情报。在英语中,“Intelligence”一词所表达的正是这种情报。目前学术界讨论较多的竞争性情报检索一词,即来源于英文中的“Competitive Intelligence”。可见,不是所有的人类社会信息都是情报。

由情报、知识、信息的定义我们可以简要总结一下三者的关系:信息是广泛存在于自然界、人类社会、人类思维之中的事物普遍联系方式;知识是人类社会中经过人类加工的系统化信息;情报尽管不全部是知识,但绝大部分的情报表现为运动着、传递着的知识。

1.2 信息素养

在信息时代,信息成为社会的重要资源。为了衡量人们在信息时代最起码应具备的信息知识和技能水平,信息素养被人们提出来以描述社会对人的信息技能的新要求。

信息素养,也译作信息素质,是一个动态的、不断完善和发展的概念。它最早是由美国信息产业协会主席保罗·泽考斯基(Paul Zurkowski)于1974年提出的。他认为,信息素养包含诸多方面:传统文化素养的延续和拓展;使受教育者达到独立自学及终身学习的水平;对信息源及信息工具的了解和运用;必须拥有的各种信息技能,如对需求的了解及确认,对所需文献或信息的确定、检索,对检索到的信息进行评估、组织和处理并做出决策等。信息素养的概念一经提出,人们对它的研究就没有停止过。