



高职高专通用系列规划教材

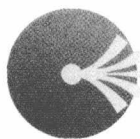
GAOZHI GAOZHUAN TONGYONG XILIE GUIHUA JIAOCAI

大学生信息检索 与论文写作

DAXUESHENG XINXI JIANSUO YU LUNWEN XIEZUO

主编 闫 瑜

HEUP 哈尔滨工程大学出版社
Harbin Engineering University Press



高职高专通用系列规划教材

GAOZHI GAOZHUAN TONGYONG XILIE GUIHUA JIAOCAI

大学生信息检索 与论文写作

主编 闫 瑜

HEUP 哈尔滨工程大学出版社
Harbin Engineering University Press

内 容 简 介

本书是在多年从事信息检索教学研究的基础上,结合高职高专院校对人才培养目标的要求以及课程自身的特点编写而成的。

全书共分十章,内容包括信息检索基础知识、网络信息资源及搜索引擎检索、报纸及其检索方法、期刊及其检索方法、会议论文及其检索方法、学位论文及其检索方法、论坛(BBS)和博客信息资源、图书及其检索方法、学术论文的撰写、论文答辩与学术规范等内容。

本书可作为高职高专院校学生学习信息检索与论文写作的教材及科技人员检索科技信息资源的指南。

图书在版编目(CIP)数据

大学生信息检索与论文写作/闫瑜主编. —哈尔滨:
哈尔滨工程大学出版社, 2010.7
ISBN 978-7-81133-816-4

I. ①大… II. ①闫… III. ①情报检索-高等学校-
教材②论文-写作-高等学校-教材 IV. ①G252.7②H152.3

中国版本图书馆 CIP 数据核字(2010)第 116012 号

出版发行 哈尔滨工程大学出版社
社 址 哈尔滨市南岗区东大直街 124 号
邮政编码 150001
发行电话 0451-82519328
传 真 0451-82519699
经 销 新华书店
印 刷 肇东粮食印刷厂
开 本 787mm×1 092mm 1/16
印 张 9.75
字 数 229 千字
版 次 2010 年 7 月第 1 版
印 次 2010 年 7 月第 1 次印刷
定 价 19.00 元
<http://press.hrbeu.edu.cn>
E-mail: heupress@hrbeu.edu.cn



随着网络化、数字化的发展,信息环境发生了很大变化,用户获取、利用信息的途径和手段也产生了根本变革,基于计算机网络的信息检索逐步成为人们获取信息的主要方式,掌握一定的信息检索、加工与利用能力,已经成为当代大学生的基本能力之一。

基于此,编者在多年从事信息检索教学研究的基础上,结合教育部颁发的《文献检索课程教学基本要求》中规定的大纲内容为准则编写了本书,可作为高职高专院校学生学习信息检索与论文写作的教材及科技人员检索科技信息资源的指南。

本书在编写的过程中,力求突出三大特色:一是实用性,以必需、够用为原则,始终把简化理论阐述、着重实际检索放在首位;二是新颖性,紧跟互联网发展步伐,以互联网上丰富的信息资源检索为主;三是通用性,学科范围涵盖文、理和工科诸多领域。

本书立足于介绍现代信息检索的方法与技术,在对信息检索基础知识进行阐述的基础上,重点介绍国内网络信息资源的检索与利用,为读者提供了较全面的信息检索知识和学术论文写作方法。全书共分信息检索基础知识、网络信息资源及搜索引擎、报纸及其检索方法、期刊及其检索方法、会议论文及其检索方法、学位论文及其检索方法、论坛(BBS)和博客信息资源、图书及其检索方法、学术论文的撰写、论文答辩与学术规范等十章内容。

本书由甘肃工业职业技术学院闫瑜副教授编写。在编写过程中,参考和引用了大量文献资料,由于篇幅所限,没能一一列出,特向有关作者致歉并表示衷心感谢!由于作者水平所限,错误和欠妥之处难免,诚望广大读者指正。

编 者

2010年1月



第一章 信息检索基础知识	1
第一节 信息的基本概念	1
第二节 信息检索基本概念	5
思考与练习题	10
第二章 网络信息资源及搜索引擎检索	11
第一节 网络信息资源基础知识	11
第二节 主要搜索引擎的使用	12
思考与练习题	31
第三章 报纸及其检索方法	32
第一节 报纸基础知识	32
第二节 报纸的检索	34
第三节 看报软件	40
思考与练习题	43
第四章 期刊及其检索方法	44
第一节 期刊基础知识	44
第二节 期刊检索数据库	52
第三节 期刊网	62
思考与练习题	63
第五章 会议论文及其检索方法	64
第一节 会议论文知识	64
第二节 会议论文数据库	65
第三节 会议网	70
思考与练习题	72
第六章 学位论文及其检索方法	73
第一节 学位论文知识	73
第二节 学位论文检索	78
思考与练习题	82
第七章 论坛、博客信息资源	83
第一节 论坛资源	83
第二节 博客资源	89
思考与练习题	92
第八章 图书及其检索方法	93
第一节 图书基础知识	93
第二节 图书分类	95
第三节 电子图书	97
思考与练习题	104



CONTENTS

第九章 学术论文的撰写	105
第一节 学术论文基础知识	105
第二节 学术论文的写作步骤	107
第三节 学术论文的基本格式	113
思考与练习题	116
第十章 论文答辩与学术规范	117
第一节 论文答辩	117
第二节 学术规范	122
思考与练习题	123
附录	124
附录 A 网上免费论文资源汇编	124
附录 B 行业信息资源汇编	134
参考文献	147



第一章 信息检索基础知识

第一节 信息的基本概念

一、信息的定义

1. 信息的概念

“信息”一词是1948年“信息论”的奠基人美国人克劳德·香农(C E Shannon)提出来的,他在研究信息理论时认为“信息是一种消息”。控制论的创始人美国人维纳(N Wiener)在他所著的《控制论和社会》一书中认为“信息是人们适应外部世界并且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内容的名称。”

自从香农和维纳共同创立了信息论之后,信息的概念立即得到普及化。近年来,“信息”一词的使用如同信息本身一样,充满了整个社会,各门学科都试图描述和说明它的本质。哲学家认为,信息是认识论的一部分;物理学家认为,信息是熵理论;数学家认为信息是概率论的发展;而有的情报学家则认为“信息就是情报”,等等。围绕信息定义的说法不下百种,除了香农与维纳的说法之外,还有“信息就是信号”,“信息就是知识”,“信息就是数据”,“信息是被反映物质的属性”,“信息是通信传输的内容”,等等。迄今为止,不同的研究学派对信息的本质及其定义还没有形成统一的意见。

在我国,尽管古往今来就有信息一词,但一直没有统一的定义标准。以前人们对信息的理解主要是指消息或某种情况的意思。我国国家标准《情报与文献工作词汇基本术语》(GB4894—85)给信息下的定义为:“信息是物质存在的一种方式、形式或运动状态,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的不确定性减少。”

信息是一个极为广泛的概念,是生物信息、自然信息、人类信息等多种信息的总称,它普遍地存在于自然界、生物界和人类社会之中。“信息”一词通常是指数据、消息所包含的内容和意义。语言、文字、符号、信号、指令、代码等都可以表达和传递信息。信息是一种独立的研究对象。我们对它的基本认识总结如下:

(1)信息不是凭空臆想的幻影,而是普遍存在的一类研究对象;它存在于自然界及人类思维之中。哪里有事物,哪里就有事物的运动,哪里就有事物运动的状态和方式。

(2)信息与事物是既有联系又有区别的两个概念。事物是信息的载体,事物的运动是信息的源泉;信息只是事物运动的状态和方式,不是事物本身。信息与事物不能等同。

(3)信息与能量也是既有联系又有区别的两个概念。传递信息需要能量,驾驭能量则需要信息,信息是事物的运动状态和方式,能量是物体做功的本领,二者之间不能画等号。

(4)本体论意义的信息是事物运动的状态和方式,认识论意义的信息是认识主体所感知或表述的事物运动状态和方式,那么人类要认识事物就必须取得信息。



2. 信息的分类

信息广泛存在于客观世界,信息是一个大家族、大系统,信息具有自己特有的结构与成分,我们要对信息进行全面系统的认识,就必须对信息进行分类。信息独立存在的形式多种多样,体现在多方面、多层次,由于人们对信息的理解与判断及需求不同,所以我们必须从不同角度按不同标准对信息进行分类,这样有助于人们全面了解与认识信息,并且促使我们理解哪些信息属于信息资源的范畴。

按信息的性质分,可分为语法信息、语义信息和语用信息。

按哲学认识主体分,可分为客观信息、主观信息。

按自然界的发展分,可分为非生物信息、生物信息和社会信息。

按信息的运动状态分,可分为半连续信息、连续信息、离散信息。

按信息产生的时间分,可分为昔时信息、现实信息、未来信息。

按信息的应用领域分,可分为工业信息、农业信息、军事信息、政治信息、科技信息、文化信息、经济信息。

按信息的逻辑意义分,可分为真实信息、虚假信息、不定信息、伪装信息。

按信息的记录与表达形式分,可分为文字信息、声像信息、数据信息、实物信息、机读信息。

按信息的生成领域分,可分为宇宙信息、自然信息、社会信息、思维信息。

按认知力与观察过程分,可分为实在信息、先验信息、实得信息。

按人的感知分,可分为直接信息、间接信息。

按信息的载体分,可分为文献信息、声像信息、光电信息、生物信息。

按信息产生效果分,可分为有用信息、无用信息、干扰信息。

按信息的内容分,可分为科学信息、技术信息、经济信息、法律信息、政务信息、金融信息、市场信息、文化信息、旅游信息。

3. 信息的表现形式

在我们身边广泛存在着信息,而我们如何来发现信息、感知信息,如何区别有用信息和无用信息呢?通常我们所发现和被感知的信息都是通过一定的媒介和它固有的传递形式所表现出来的。

(1)文字信息 文字信息是人类便于信息的交流、联系所独创的一种形象符号。它包括 ASCII 码、电报码、单元代码、二进制码、各类字符符号,这些代码、符号、文字字符等都是信息的表现形式。

(2)图像信息 图像信息以图像、图形为主要表现形式,图像信息较于文字信息更直观,包括人工图像信息和自然景象信息。人工图像信息指人类通过某种制图软件和工具制作的图形和图像等从视角上所反映出来的信息。自然景象指在自然界中客观表现出的某种间接或是抽象的图形图像信息。

(3)数据信息 数据信息是以数据为主要表现形式的信息,数据信息包括人类统计的数据、测量的数据、通过某次实验或某种仪器表现出来的数据等。

(4)语音信息 语音信息主要以声音为表现形式,包括人们日常的演讲、交流、演奏等所产生的信息。

二、信息的内涵

1. 心理学领域对信息的认识



从心理学角度对信息现象进行研究认为信息是存在于人的意识之外的东西,是客观存在的,它不依附于人的意志而存在。

2. 通信与信息科学领域对信息的认识

通信理论的重要研究成果之一就是信息科学的出现。我国学者钟义信在1988年出版的《信息科学原理》一书中认为,信息是事物运动的状态与方式,是物质的一种属性。信息不同于消息,消息只是信息的外壳,信息则是消息的内核;信息不同于信号,信号是信息的载体,信息则是信号所载荷的内容;信息不同于数据,数据是记录信息的一种形式,同样的信息也可以用文字或图像来表述;信息不同于情报,情报通常是指秘密的、专门的、新颖的一类信息,可以说所有的情报都是信息,但不能说所有的信息都是情报。

3. 经济学领域对信息的认识

从经济学角度对信息进行研究,在各经济学家眼中,信息、物质、能量是客观世界并列的三大要素,信息是为管理和决策提供依据的有效数据。

4. 哲学领域对信息的认识

在哲学研究领域中,哲学家们通常从产生信息的客体对信息进行定义,他们认为事物的特征通过一定的媒介或传递形式使其他事物感知。这些被感知的、表现该事物特征的信号是该事物向其他事物传递的信息。这一定义说明信息是事物本质、特征与其运动规律的反映。不同的事物有不同的本质、特征和运动规律,人们通过事物发出的信息来认识或区别于其他事物。

5. 新闻领域对信息的认识

在新闻界,信息普遍被理解为是对事物运动状态的表述,是人—人、物—人、物—物之间的特征传输,通常人们认为新闻是信息中的一种。

6. 信息管理领域对信息的认识

在信息管理领域中对信息的理解与认识受计算机领域信息技术发展的影响。美国学者史密斯(A N Smith)和梅德利(D B Medley)在《信息资源管理》中认为,信息是数据处理最终产品,是经过收集、记录、处理,以检索的形式储存的事实或数据。

7. 图书馆学和情报学领域对信息的认识

图书情报界将信息定义为事物或记录,而记录则包含文字与图片,这些文字和图片信息是相对读者而言的。相对来说,文字和图片信息不能脱离读者而独立存在,读者通过阅读和记录理解这些信息,信息与读者的认识结构是相对应的。

导致以上对信息不同定义的主要原因有:首先,信息本身是一个多元化、多层次、多功能的复杂综合物;其次,信息科学是一门新学科、大学科,它有许多分支学科,其内涵与外延伴随着社会、经济与科学技术不断地发展;第三,人们对信息根据不同的研究与使用目的,从不同的角度与层次着手,对信息的理解与解释也不尽相同。由于这些主要原因的存在,导致对信息的定义产生差异性、多样性。随着人类的进步和科学技术的不断发展,人们对信息的定义不断深化与明确,并且对自然界和社会的各类信息通过获取、识别来区分各类事物。人们可以从不同层次、不同的侧面对信息的概念给予不同的解释,对信息赋予不同的内涵与外延,从而达到认识世界与改造世界的目的。

三、信息的性质与特征

信息的定义揭示了信息的本质属性,信息来源于物质,而又不是物质本身;信息来源于

精神世界,但也不限于精神领域;信息最终归结于物质的普遍性,是物质的运动状态与运动方式。信息的物质属性决定了信息的一般属性特征,这些属性特征包括客观性、普遍性、依附性、可传递性、可储性、可转换性、可伪性、时效性、动态不定性、差异性、可预测性、共享性等。

1. 客观性

任何物质都是信息的母体,信息是对物质存在的状态、性质与特征变化的客观反映,而这种反映物质状态、性质与特征的信息是客观存在的,这种客观存在的信息可以被人类感知、获取、整理、传递和利用。

2. 普遍性

不管是自然界还是人类社会,每时每刻都在产生信息。只要有物质存在,就有物质运动,有物质运动就有信息产生,信息产生是无限的,这就决定了信息无处不在,无时不有,物质运动的普遍性决定了信息的普遍存在性。

3. 依附性

信息本身看不见、摸不着,也不能独立存在和自动交流。信息只有借助于某一载体或媒介才能被交流和利用、继承和传播共享,信息不可能存在于任何物质之外。信息可以通过文字、语言、符号、代码等方式依附在纸张、胶片、磁盘、光盘等不同载体上,其内容并不因为记录方式或载体不同而发生改变。信息因依附这些载体而变成人类的一种广泛的信息资源与财富。

4. 可传递性

信息的价值主要体现在发挥和利用上,而信息的发挥与利用则体现在信息的传递上。信息的传递可将信息从一个时段传递到另一个时段,从一个地方传递到另一个地方。信息的传递分为时间传递和空间传递,信息的存储属于时间上的传递活动,而空间上的传递活动就是我们所说的通信。因为信息具有可传递性才使得人们可以共享信息,在共享的基础上才能使信息发挥更大的作用。

5. 可储性

信息是具体存在的,它不是可以任意想象或随意创造的物质,它的可知性使信息可以通过记忆、收集、整序、综合、加工等方式进行存储。大脑对信息的存储就是我们通常所说的记忆。人类对信息的存储主要表现在载体上,诸如图书、期刊、档案、资料、光盘、磁盘、磁带等介质可以存储大量的信息,信息通过有序存储可以实现信息的累积。人们对信息进行存储与积累是为了以后能方便、快捷和有效地利用信息。

6. 可转换性

人类为了更好地对信息进行转换、加工处理,发明创造了一系列的相关仪器、技术和方法,运用这些仪器、技术和方法对信息进行编码、压缩和有序化转换等操作,使信息以不同形态和方式来表达与转换,例如同一信息内容既可用语言文字来表示,也可用电流、声波和光波等来表示;同类信息可依附于纸张、胶片、光盘、磁盘等不同的物质载体,而这些物质载体之间可以相互进行转换,并且其信息内容保持不变。信息的可转换性也可称为信息有序性,对信息进行一系列的加工整理,使无序信息变成有序信息,有序信息便于流通、传播与利用,同时有利于信息继承与再生。

7. 可伪性

信息的可伪性包括两个方面,一方面,因种种原因,人类借助于仪器或载体使人类在感知、接收、识别物质信息过程中产生对客观事物的错误认识与判断;另一方面,信息的相对独



立性使信息脱离原事物后,人们凭主观认识来理解信息,由于人们的认识能力或动机不同,存在对信息的片面性认识,诸如“盲人摸象”所产生的虚假信息,又称为伪信息。

8. 时效性

物质世界中事物是不断变化的,而信息则跟随事物的变化而不断变化,无论是针对信息整体还是某条具体的信息,它的信息内容都会随着时间的变化而变化。当人们将特定时间的事物运动的信息通过输入、处理、传递、输出等提取过程时,由于信息的提取与信息本身状态的时间差和事物的运动规律致使信息的效用逐步减弱,以至失去价值。信息价值的体现一般分为四个阶段:升值期,峰值期,减值期和负值期。信息在不同周期呈现不同的价值,在现代信息社会,信息的应用周期越来越短,如不及时把握与利用最新信息,信息就会贬值,只有充分重视和发挥信息的时效性,才能将信息转化为时间与金钱。这表明了信息的时间价值与信息经济价值,这就是所谓的信息的时效性。

9. 动态不定性

信息是动态不定的,信息所反映出的总是在特定时刻事物运动的状态和方式,当人们将该时刻的信息提取之后,事物仍然在不停地运动。古希腊哲学家曾言:“人不能两次踏进同一条河流”,这就是所谓信息动态不定性的形象说明。

10. 差异性

信息相对用户的使用存在着差异性,由于人们对信息的需求、理解和判断力有别,所以相同的信息对不同的使用者有着不同的作用,同时信息的价值体现也不同。

11. 可预测性

人们利用成熟的理论和相关的仪器设备,在掌握大量信息的基础上,通过有计划、有目的的预测、推理、调查等方法来获得未来的发展,未来的发展信息一般超前于现实信息,同时我们又称为信息的超前性。

12. 共享性

信息与物质和能量的本质区别就是信息具有物质和能量不具备的共享性。一般来讲,物质和能量归所有者拥有,在互相交换过程中所有权也随着交换,一方失去,一方获得。而信息却不同,同一内容的信息可在同一时间不同地点由两个或两个以上的多个用户获得使用,并且在信息的传递与交换过程中,原信息提供者并不会因为提供了信息而失去原有的信息内容和信息量。换句话说,就是不会因为一个人得到了信息,而另一个人就失去信息。信息可以无限制地进行复制与传播,可以分配给大家所共享,信息不会因为多人重复复制与使用而减少,相反,信息的广泛交互式交流可产生更多的信息。英国文学家萧伯纳有一形象的比喻:倘若你有一个苹果,我也有一个苹果,我们彼此交换,我们双方仍然各有一个苹果。但是,倘若你有一种思想,我也有一种思想,我们彼此交流这些思想,我们两个人各有两种思想。在这里所说的思想,我们可以理解为信息,信息的共享性可以提高信息的利用率和信息的容量。

第二节 信息检索基本概念

一、信息检索的含义

信息检索(Information Retrieval)的概念有狭义和广义之分。狭义的信息检索是指依据一

定的方法,从已经组织好的大量有关文献集合中,查找并获取特定的相关文献的过程。这里的文献集合,不是通常所指的文献本身,而是关于文献的信息或文献的线索。如果真正要获取文献中所记录的信息,那么还要依据检索所取得的文献线索索取原文。

广义的信息检索包括信息的存储和检索(Storage and Retrieval)两个过程。信息存储包括三个步骤:

(1)信息的选择与收集 它是指检索系统根据本系统的服务目的,确定信息收集、处理的原则,对分布在各处的离散信息进行收集加工;

(2)信息的标引 标引是信息加工人员对收集到的信息内容特征进行分析之后,对每条信息加上系统能够识别的检索标志的过程;

(3)形成大量有序可检的信息集合 工作人员将标引后的信息条目录入,并将其按照一定的顺序排列起来,形成有序的信息集合——数据库,从而为信息检索奠定基础。

信息的检索过程则是信息存储的逆过程。信息用户根据自己的需求对主题和概念进行认真分析后,将自己的信息需求转化为检索表达式,该检索表达式与系统标志的比较匹配过程就是检索的过程。

二、信息检索的类型

文献信息检索是指从汇集的大量文献中迅速而准确地查找出特定用户所需的信息,按不同的标准可以将其划分为不同的类型。

(1)根据检索对象的不同可以分为文献检索、事实检索、数据检索。

(2)根据信息检索结果的内容可以分为书目检索、数据检索、全文检索、事实检索、声像检索。

(3)按照信息检索技术的发展可划分为全文文本检索、多媒体检索、超文本检索、网络信息检索、资源检索。

①文献检索(Document Retrieval) 文献检索是以查找文献线索为对象的检索。从文献中查出关于某一主题、著者、机构、地域、时代的有关信息,查找某篇论文出处以及某一出版物的收藏处等均属文献检索范畴。

文献检索只是一种相关性的检索,它提供的只是文献线索(题录或文摘),一般需要阅读文摘后才能决定取舍。它主要是通过二次文献(检索工具如目录、题录、文摘等)进行检索。

②事实检索(Fact Retrieval) 事实检索是以查找数值与非数值混合情报为对象的检索,或者说是以事实信息为检索对象的检索,包括事物的性质、定义、原理以及发生的地点、时间、前因后果等。例如,了解“IBM 公司 1998 年推出了哪些新型计算机?”“中国有几家公司进入了世界 500 强?”等。

事实检索是一种确定性检索,但是检索到的结果必须经过分析、对比、研究之后才能应用。

③数据检索(Data Retrieval) 数据检索是以查找数据、公式或图表为对象的检索。例如,查找科技数据、金融数据、人口统计数据、公式、图表、价格、某种物质的分子式、某种设备的型号与技术参数等,均属数据检索。

数据检索也是一种确定性检索,其结果是肯定的。数据检索主要利用各种字典、词典、百科全书、年鉴、手册、名录等参考工具书,也可以通过计算机查找网络上大量的动态数据。

④书目检索 书目检索指查出某一主题的文献条目的检索,按检索结果可分为题录检索、文摘检索、图书与期刊等目录检索。



⑤声像检索 声像检索主要是指对声音、图像和其他图文信息的检索。

⑥全文检索 全文检索是指利用原始文献库进行的检索,检索的结果可以是原始文献的全文,也可以是其部分内容。

⑦全文文本检索 全文文本检索也叫做全文数据库检索。它通过计算机将文件的全貌,包括文字和图形、图像等非文字信息转换为计算机可读形式,我们称之为“记录”。计算机索引程序对每一条记录按词建立一个索引,指明该记录在文章中出现的次数和位置。当用户查询时,检索程序就根据事先建立的索引进行查找,并将查找的结果反馈给用户的检索方式。这个过程类似于通过字典中的检索字表查字的过程。

⑧超文本检索 超文本检索是一种具有联想式思维功能的新型检索技术。它与传统检索方式不同的是,后者的检索是一个单向的孤立过程,而超文本检索系统提供的是一种多向的网络检索环境。检索时,用户能够从任何一个节点开始,从不同角度检索到其感兴趣的信息。

⑨多媒体检索 能支持两种以上媒体的数据库检索就称之为多媒体检索。检索集文字、图形、图像、动画、声音为一体的多媒体数据时,不仅能够浏览查询对象的文字描述,而且还能做到听其声、观其形。

⑩网络信息资源检索 网络信息检索系统是一个集各种新型技术于一体的,能够对各种类型、各种媒体的信息进行跨时间、跨地区检索的大系统。网络信息资源的组织管理需要诸多信息技术的支持,其中以 WWW 全球浏览技术最具优越性和可用性。WWW 是一种集超文本技术、多媒体技术和网络技术于一体的新型检索工具。

三、信息检索方法

1. 常用法

常用法又称直接法,是指直接利用检索工具(系统)检索文献信息的方法,这是文献检索中最常用的一种方法。它又分为顺查法、倒查法和抽查法。

(1)顺查法 顺查法是指按照时间的顺序,由远及近地利用检索系统进行文献信息检索的方法。这种方法能收集到某一课题的系统文献,适用于较大课题的文献检索。例如,已知某课题的起始年代,现在需要了解其发展的全过程,就可以用顺查法从最初的年代开始,逐渐向近期查找。顺查法的优点是漏检、误检率低,缺点是劳动量大。

(2)倒查法 倒查法是由近及远,从新到旧,逆着时间的顺序利用检索工具进行文献检索的方法。这种方法的重点在近期文献上,因此可以很快地获得最新资料。例如,进行新课题立项前的调研就可用此法。使用这种方法劳动量虽小,却容易造成漏检。

(3)抽查法 抽查法是指针对检索课题的特点,选择有关该课题的文献信息最可能出现或最多出现的时间段,利用检索工具进行重点检索的方法。它适于检索某一领域研究高潮很明显的、某一学科的发展阶段很清晰的、某一事物出现频率在某一阶段很突出的课题。这是一种检准率较高又比较省时的方法。

2. 追溯法

追溯法是指利用已经掌握的文献末尾所列的参考文献,进行逐一地追溯查找引文的一种简便的扩大情报来源的方法。它还可以从查到的引文中再追溯查找引文,像滚雪球一样,依据文献间的引用关系,获得越来越多的相关文献。

3. 综合法

综合法又称为循环法,它是把常用法和追溯法加以综合运用的方法,也是实践中采用较多的一种方法,它兼有常用法和追溯法的优点,可以查得较为全面而准确的文献。

四、信息检索途径

利用各种索引语言,对文献的外部特征和内部特征进行描述,其结果就产生不同的文献标志。大量的文献标志,按照字顺的次序或逻辑的次序排列起来,就产生了系列化的、可供检索的文献描述体(书目文档)——这就是各种类型的索引。信息用户在检索时,把所需信息的某种特征标志转换为文献标志,以此为入口进行检索,这个检索入口就叫做检索途径,多种多样的索引就可以提供多种多样的检索途径。

一般来讲,常用的、重要的检索主要有以下几种途径。

1. 题名检索途径

题名途径依据的是检索工具中的书名索引。属于这一索引系统的有书名目录(索引)、刊名索引、篇名索引、标准名称索引、数据库名称索引等,这些可统称为题名索引。题名索引主要在计算机检索系统中应用较多,我国古书目录最重要的检索途径也是题名检索途径。

2. 著者检索途径

著者检索途径是指根据已知文献著者来查找文献的途径,它依据的是著者索引,包括个人著者索引和机关团体索引。

3. 分类检索途径

分类检索途径是按照文献资料所属学科(专业)类别进行检索的途径,能够满足信息用户族性检索的要求。所依据的检索工具就是分类索引。

分类检索途径有利于从学科或专业角度广泛地获得较系统的文献信息,达到较高的查全率。分类途径检索的关键在于正确理解检索工具的分类表,使用者必须了解学科体系。

4. 主题检索途径

主题检索途径是一种主要的检索途径,它通过文献资料的内容主题进行检索。主题检索途径依据的是各种主题索引或关键词索引。在国外现行的排检方法中,主题法是最基本的方法,因而在检索国外文献信息时,主题检索是最常用的。主题检索途径的关键就在于分析项目,提炼主题概念,运用词语来表达主题概念。

五、信息检索的步骤

信息检索工作是一项实践性和经验性很强的工作,检索程序与检索的具体要求有密切关系,对于不同的项目,可能采取不同的检索方法和程序。文献信息检索大致可分为以下四个步骤。

1. 分析研究课题,明确检索要求、时间、范围

分析研究课题的目的就在于明确课题所要解决的问题,把握关键,有的放矢,这是检索效率高低或成败的关键。

(1)首先应分析课题的内容实质、所涉及的学科范围及其相互关系,明确所要检索的文献内容、性质等,根据检索课题抽提出能准确反映课题核心内容的主题概念,明确主要概念与次要概念,并初步确定出逻辑组配。

(2)根据检索课题的检索目的和要求,确定检索年限、语种、文献类型等。



①确定检索年限主要根据研究课题的背景信息,如起始年代和研究的高峰期等。一般来说,检索的时间范围应根据检索课题的具体情景而定。例如,进行查新检索时,就需要检索最近10年的文献;若是纯属掌握动态或解决某一个具体问题,则以满足需要为准,时间可长可短。

②检索语种的范围主要是依据课题的检索范围。确定文献类型时应在主题分析的基础上,根据检索目的和要求,明确课题对检索深度的要求。例如,课题属于探讨基础理论性的,则所检索的文献类型应以期刊论文、会议文献的一次文献为主;如果课题属尖端科技,则应侧重于查科技报告;如果需要查新,则应以检索专利文献为主。

(3)分析用户对检索的评价指标是查新、查全还是查准;用户需要提供的是题录、文摘还是全文?根据用户的检索要求,进行归纳课题已知的检索线索,如专业名词、术语、分类号、主题词、著者姓名等,为下一步检索实践提供准确可靠的依据。

2. 确定检索策略

检索策略的好坏,直接影响到相关文献的查全率和查准率,关系到检索的效果。检索策略具体地说,就是检索提问逻辑表达式,就是在分析用户信息实质的基础上,确定检索途径与检索用词,并明确各主题词之间的逻辑关系与查找步骤的安排。

制定检索策略,优化检索过程,主要涉及四个方面的问题。

(1)选择检索工具 选择恰当的检索工具,是成功实施检索的关键。选择检索工具一定要根据检索课题的内容、性质来确定,主要应从以下几个方面来考虑:

①从内容上考虑检索工具的学科专业范围。选择检索工具时,应以专业性检索工具为主,综合型检索工具进行配合、补充。在技术和手段上,由于计算机检索系统适应多点、多属性的检索,检索精度高,应首选机检工具,而且应选择合适的数据库。

②如果只有手工检索工具,应选择专业对口、文种熟悉、索引体系完善、报道及时、揭示文献信息准确、有一定深度的手工检索工具。如果一种检索工具同时具有机读数据库和印刷型文献两种形式,应以检索数据库为主,这样不仅可以提高检索效率,而且还能提高查准率和查全率。

为避免检索工具在编辑出版过程中的滞后性,在必要时应补充查找若干主要相关期刊的现刊,以防漏检。

(2)确定检索途径 检索工具确定后,就需要根据具体的检索工具来确定检索点,即检索途径。一般的检索工具都根据文献的内部特征和外部特征提供多种检索途径。各检索途径都有各自的特点和长处,选用何种检索途径,应根据检索课题的要求及所包含的检索标志和检索系统所提供的检索途径来确定。当检索课题内容涉及面广,文献需求范围较宽,泛指性较强时,宜选用分类途径;当课题内容较窄,文献需求的专指性较强时,宜选用主题途径;当只知道物质分子式时,宜选用分子式途径;当选用的检索系统提供的检索途径较多时,各种检索应综合应用,互相补充,避免单一途径不足所造成的漏检。

(3)优选检索方法 优选检索方法的目的在于寻求一种快速、准确、全面地获得文献信息的检索效果。

(4)制定、调整检索策略 检索工具、检索途径、检索方法确定后,就需要制定一种可执行的方案。手工检索系统的检索策略由于检索系统的限制,每次检索只能从一个检索点出发,因此也就只能一边检索一边分析取舍,从而获得用户所需要的文献信息。在计算机检索的条件下,由于信息提问与文献标志之间的匹配工作是计算机进行的,因此构造精确的检索式是执行计算机检索的前提,它可将各检索单元之间的逻辑关系、位置关系等用检索系统规

定的各种组配符连接起来,实施有效检索。但这个检索式不是一成不变的,要把检索结果与用户需求不断地进行判断、比较之后,对检索式进行相应的修改和调整。

3. 查找文献线索

在明确了检索要求,确定了检索系统,选定了检索方法,就可以应用检索工具实施检索,所获得的检索结果即为文献线索,对文献线索的整理、分析、识别是检索过程中极其重要的一个环节,需要作好以下两个方面的工作:

(1)作好检索记录。做好检索记录的目的在于必要时进行有效核对。包括记录好使用检索工具的名称、年、卷、期、文献号(索引号)、文献题名(书名)、著者姓名及其工作单位、文献出处等。检索工具中,文摘、题录所著录的文献出处,是索取原始文献的关键因素。

(2)在检索工具中,文献出处项对摘录的文献类型不加明显区分,需由检索者自己进行辨别。只有识别出文献类型,才能确定该文献可能收藏在何处,查何种馆藏目录,如何借阅和复制。

4. 索取原始文献信息

信息检索的最终目的就是获取原始文献。当检索到文献线索并识别文献类型以后,即可根据不同的文献类型和语种索取原始文献。传统的原文索取方法是根据检索到的文献线索,再利用馆藏目录查找收藏单位、收藏点,采取借阅或复制等方式索取原始文献。但是,随着网络技术的飞速发展,全文数据库的兴起,使得原始文献信息的获取方式多种多样。归纳起来,原始文献有如下几种索取方法:

(1)向著者索取原始文献 根据文献线索所提供的著者姓名及其工作单位等可直接与作者联系,索取原始文献;

(2)利用馆藏目录、公共查询系统、联合目录获取原始文献 查到本馆文献检索者可利用馆藏目录,但是独立的一个馆,其馆藏毕竟有限,读者需要的文献若是本馆没有收藏的,就需要借助 OPAC 和联合目录实施馆际互借,或者通过文献传递获取;

(3)利用网上全文数据库获取原始文献 现在有许多全文数据库可以为用户提供直接检索。提供中文期刊全文的数据库如“维普中文科技期刊数据库”,“中国期刊全文数据库”,“万方数字化期刊”等;提供中文图书全文的数据库如“书生之家”,“超星数字图书馆”,“方正 Apabi”等。

思考与练习题

- 1-1 简述信息的概念及其分类。
- 1-2 简述信息的基本特征。
- 1-3 信息检索的常用方法有哪些?



第二章 网络信息资源及搜索引擎检索

第一节 网络信息资源基础知识

一、网络信息的概念

网络也就是人们所称的计算机网络,是指将分散在各处,却具有独立功能的多台计算机终端及其附属设备,通过通信设备和线路连接起来,运用功能完善的通信软件按照网络协议进行数据通信,以实现资源共享的系统。

人们在网上全方位的浏览与交流使网络信息资源应运而生。虽然网络信息资源一词已被越来越多的人所接受,但至今仍没有一个统一的定义。较为流行的两种表述方式是:其一,网络信息资源就是通过计算机网络可以利用的各种信息资源的总和;其二,网络信息资源是指以数字形式记录的,以多媒体形式表达的,存储在网络和计算机磁介质、光介质以及各类通信介质上的,并通过计算机网络通信方式进行传递的信息内容的集合。我们在理解这一概念时,需要特别注意的是网络信息资源并非包含所有投放到因特网上的信息,而只是指其中能满足人们信息需求的那一部分。

二、网络信息的类型

网络信息资源的类型根据不同的标准,可以划分不同的类型。人们习惯于按照信息交流的方式来分类,分为正式出版信息、半正式出版信息和非正式出版信息三种类型。

1. 正式出版信息

指受到一定的产权保护、信息质量可靠、利用率较高的知识性、分析性信息。如通过万维网,用户可以查询到的各种数据库、联机杂志和电子杂志、电子工具书、报纸、专利信息等。

2. 半正式出版信息

指受到一定产权保护的但没有纳入正式出版信息系统中的信息,如以各种学术团体和教育机构、企业和商业部门、国际组织和政府机构、行业协会等单位的网址或主页上,我们可以查询到从正式的出版物系统中无法得到的“灰色”信息。

3. 非正式出版信息

指流动性、随意性较强的,数量大,质量难以保证和控制的动态性信息,如电子邮件、电子会议等。

三、网络信息的特点

同其他媒体信息相比,互联网是世界上最大的信息宝库,不利用互联网信息资源,就很难达到相关领域的前沿,网络信息资源具有诸多鲜明的特征。

1. 内容无限

网络是一个庞大的信息资源资料库,而且每天都在增加。信息资源极为丰富,覆盖面