

信息检索与网络基础

主编 张承华

INFORMATION—RETRIEVALING AND INTERNET FOUNDATION



山东文化音像出版社

信息检索与网络基础

主编 张承华

山东文艺音像出版社

新出音管〔2000〕547号

信息检索与网络基础

Xinxi JianSuo Yu WangLuo JiChu

主 编 张承华
责任编辑 陈留彬
封面设计 张卫锋
出版发行 山东文化音像出版社
出版社地址 济南市山大北路27号 邮编 250100
经 销 新华书店
印 刷 山东工人报印刷厂
开 本 850×1168 毫米 1/32
印 张 9.75
字 数 250千字
版 次 2001年8月第1版 2001年8月第1次印刷
印 数 1~2500册
书 号 ISBN 7-980047-84-2/TP·23
定 价 18.80元

版权所有 盗版必究

主 编	张承华			
副主编	张 伟	刘彦辉	王春莉	唐仕善
	马方亮	范翠萍	刘士新	张树新
编委会				
成 员	马方亮	王春莉	刘彦辉	刘士新
	唐仕善	张承华	张 伟	张树新
	范翠萍	李君之	孙守军	
主 审	杨荫楼			

内容提要

本书针对当前信息检索与利用教学 and 实际工作的需要,系统地阐述了信息检索的基本原理和基本方法,介绍了国内外一些重要的信息检索工具和使用方法。本书比较全面地论述了文献信息的产生和发展过程,对论文写作的步骤、格式、注意的问题以及计算机检索和国际互连网络等方面的知识都作了较详细的介绍,并给出了较多的检索实例,供读者在学习中参考或在实际工作中选用。

本书反映了当前典型检索系统的最新进展,内容体系上有新的突破,在内容上适当压缩了手检比例,适当扩充了机检知识,尤其是增加了因特网信息利用方面的内容,介绍了科研的一般程序和论文写作方法,读后使人有一个全新的感觉。

本书可作为各类学校信息检索课的教材,也可供科研人员、教师和工程技术人员作为信息库入门的钥匙,亦可供图书馆专业人员参考。

前 言

知识经济是以知识的创新及其应用为主要发展动力的经济形态,而信息产业是推动知识经济发展的主要动力。一个国家能否在全球性的知识经济时代拥有较强的国际竞争和可持续发展能力,除了加大在建设知识与技术创新体系方面的投资和政策力度之外,还必须有一个高度信息化和社会化的知识生产、传播、扩散的基础设施,来保证人们在知识信息生产、传播、扩散和应用的各个过程中进行有效的社会化合作。在知识经济条件下,具有强烈的信息意识并能迅速准确地获取和利用各种信息资源,已经成为各级各类人才在现代社会获得成功的基本素质。在当今的信息时代,人们在社会中无论从事何种工作,能够从数量庞大、瞬息万变、类型复杂、文种多样、出版分散的信息群中,迅速、准确、完整、经济地获取与自己研究课题相关的文献信息资料,是必备的科学劳动能力之一。任何一位劳动者,在他从事创造性的研究活动时,在他运用观察、比较和推理的能力来研究自然现象和社会现象时,必须首先详细地占有资料,在前人已经取得成就的基础上去进行新的探索。俄罗斯文献信息学家布留索夫说:学问与其说是知识的储蓄,倒不如说是善于在浩如烟海的信息群中找到知识的本领。科研的过程,在很大程度上是从搜集文献信息入手,经过分析研究到引出科学结论的过程。其中的每一个步骤,都离不开信息的检索与利用。系统地阅读有关文献是掌握研究课题的历史、现状及其发展趋势的重要途径。历史上有成就的学者能够在科学领域取得辉煌成就,无一不是在搜集和积累文献信息上下过苦功夫。

一项权威的调查表明,我国科技工作者和各级各类学校的师生获取文献信息的能力较低,大多数不善于利用图书馆,不具备信

息检索与利用的能力,不会利用书目、索引、文摘以及其他书目参考工具查找所需文献信息,对信息港、国际互联网络等先进的信息处理和传递方式感到难以适应。因此,对读者普及信息检索与利用知识具有明显的重要性和紧迫性。

1984年至1998年原国家教委多次发文,要求各级各类学校普遍开设文献信息检索与利用课,这是一项具有战略意义的决策,不仅有助于教育质量的提高,而且是教育面向未来的一个组成部分,是全面提高学生素质,培养跨世纪合格建设人才不可缺少的一个重要教育环节。对于学生不断吸取新知识,改善知识结构,提高自学能力和研究能力,发挥创造才能都具有重要意义。有鉴于此,我们编著了此书,力求能深入浅出地介绍有关信息检索与利用方面的知识,目的在于培养学生不断提高获取信息的能力,从而增强毕业生的职业适应能力和创新能力。

本书由张承华主编,杨荫楼主审。张伟、刘彦辉、王春莉、唐仕善、马方亮、范翠萍、刘士新、张树新为副主编。参加本书编写的有王春莉、马方亮、刘彦辉、刘士新、唐仕善、张承华、张伟、张树新、范翠萍、李君之、孙守军。

本书在组稿和编审过程中,参考了许多同志编著的相关文献资料,听取了许多同志提出的宝贵意见,由于篇幅所限,不便于全部标出,在此一并表示感谢!信息检索的理论和实践发展变化很快,一些新的理论、方法层出不穷。本书的编写,由于时间仓促,而编著者学识水平有限,错误难免,尚望广大读者不吝赐教,至为感激。

张承华

2001年6月于泉城千佛山

目 录

第一章 信息检索概论	(1)
第一节 信息检索的作用	(1)
第二节 信息检索的研究对象和类型	(5)
第三节 文献的功用与类型	(9)
第四节 现代文献信息的发展特点及趋势	(18)
第二章 信息检索原理	(26)
第一节 文献信息检索原理	(26)
第二节 检索语言	(28)
第三章 信息的收集和检索	(41)
第一节 信息收集的渠道和方法	(41)
第二节 文献信息检索的途径与方法	(49)
第三节 检索工具	(56)
第四章 中文检索工具	(66)
第一节 中文检索工具概述	(66)
第二节 常用的中文检索工具	(67)
第五章 国外科技检索工具	(76)
第一节 美国《工程索引》Ei	(76)
第二节 美国《化学文摘》CA	(81)
第三节 英国《科学文摘》SA	(85)
第四节 日本《科学技术文献速报》	(91)
第五节 美国《科学引文索引》SCI	(95)
第六节 其它国外文献检索工具	(103)
第六章 特种文献检索	(109)

第一节 专利文献检索	(109)
第二节 标准文献检索	(117)
第三节 学位论文及其检索	(123)
第四节 会议文献检索	(128)
第五节 科技报告及其检索	(137)
第七章 参考工具书	(145)
第一节 参考工具书概述	(145)
第二节 综合性参考工具书选介	(154)
第三节 事实检索和数据检索	(161)
第八章 计算机信息检索	(175)
第一节 计算机信息检索概述	(175)
第二节 计算机信息检索原理	(180)
第九章 联机信息检索	(190)
第一节 联机检索概况	(190)
第二节 DIALOG 系统联机检索方法	(195)
第三节 STN 系统联机检索方法	(204)
第十章 光盘数据库检索	(208)
第一节 光盘检索系统概述	(208)
第二节 中国学术期刊(光盘版)检索	(211)
第三节 DIALOG 系统光盘数据库检索	(225)
第四节 UMI 公司光盘数据库检索	(228)
第五节 其它光盘数据库简介	(230)
第十一章 因特网信息检索	(244)
第一节 因特网的产生与发展	(244)
第二节 因特网的功能与特点	(249)
第三节 因特网的连接	(255)
第四节 因特网信息资源的利用	(257)
第十二章 科技论文写作	(274)

第一节 科研课题的设计与实施	(274)
第二节 科技论文写作	(280)

第一章 信息检索概论

第一节 信息检索的作用

信息是事物存在和运动状态及其特征的反映。人类发展的历史就是在获得信息,认识信息,传递信息,利用信息和创造信息的进程中,通过信息来认识世界和改造世界的。随着计算机技术和数据传输技术的迅速发展,人类的经济活动正在演变为知识经济,世界全面进入了信息时代,信息的集中地即财富和权力的集中地已成为共识。在人类社会中谁占据信息制高点,谁就占有了社会财富已成为不争的事实。

人类进入 20 世纪以来,科学技术以前所未有的高速度向前发展,研究规模迅速扩大,各种研究成果大量涌现,导致信息量急剧增加。科学研究不再局限于“教授的小试验室中或在发明家后院的小书房中”。从 42 万科学家参与的阿波罗登月计划到耗资数千亿美元,近百万科学家参与的全球信息高速公路,它的每一项发明创造,都具有连续性和继承性的特点,都是信息积累和转换的结果。任何一个科技人员,在开始研究某一项课题之前,必须掌握相关的科技信息,有赖于在前人已经取得成就的基础上进行新的探索。

面对浩如烟海的信息群,迅速、准确、没有重大遗漏地检索出与科研创新和工作生活相关的信息资料,就要掌握信息检索的方法。英国信息学家简休森和约翰逊曾分别提出,人类的知识应分为两部分,一部分记忆在脑海里,另一部分只要知道如何查找就可以了。古人说得好“授人以鱼,只供一饭之需,教人以渔,则终身受用无穷。”如果把有用的信息比作“鱼”,掌握了信息检索与利用的技

能,就是学会了“捕鱼法”。在科技日新月异发展的今天,信息检索的意义和作用更加突出,具体表现在以下几个方面:

一、信息检索是知识更新的手段

一个成熟的科技工作者,学校教育传授给他的知识仅占 10%~20%,其余的知识则是走向社会后通过获取信息而充实的。学校教育仅仅是一个阶段性的教育。科技工作者必须依靠检索能力,学会获取知识、更新知识的手段,学会不断进取的自学方法。

二、节省查找信息的时间

信息量的过分庞大和迅速增长,加重了科研人员搜集信息的负担。据科学院一位研究员介绍,50年前,搞一个研究课题,只须到图书馆查两周资料,即可动手做实验了,现在花费大量时间在图书馆查找,也掌握不了课题的全面情况。德国教授柏英浩统计,现在一位科学家即便是夜以继日地阅读,也只能涉猎世界上有关书刊的 5%。而学会正确的信息检索方法,使用光盘或联机检索,一般只须几分钟即可检全所需文献信息。一般情况下,科技人员查找信息资料所花费的时间约占整个课题研究时间的 1/3 左右。如果有完善的检索设施和科学的检索方法,无疑会节省研究人员的大量时间,腾出更多的精力从事创造性的劳动,提高科研效率。

三、管理工作正确决策的依据

准确、及时、可靠的信息是科学决策的基础,而信息的获得则依赖于检索策略的实施。我国前几年由于信息不灵,盲目重复引进,使一些引进技术和设备不适用,造成了惊人的损失和浪费。例如我国河南省某药厂从瑞士引进了不成熟的维生素 C 生产技术,损失人民币十几亿,未及投产即宣告破产。与此相反,我国在引进英国皮尔金顿公司的浮法玻璃生产线的过程中,对方提出技术转

让费为 2500 万英镑(约合人民币 2.5 亿元)。经过中方进行专利文献检索查明,该生产线所含有的 137 项专利技术中有 51 项已经失效,成为公共技术。中方据此检索结果迫使英方将转让费降为 52.5 万英镑,节约近 2 亿元人民币。随着市场经济的确立,经营者在技术引进、立项、成果鉴定、专利申请、市场调研、实施技术保护、市场保护、人才保护、经济纠纷索赔等一系列事务中,都离不开信息检索,若利用得法,可避免重大的损失和浪费,获得最佳经济效益和社会效益。

四、继承和借鉴前人成果,避免重复研究或走弯路

科学巨人牛顿说过:假如我比别人略微看得远些,那是因为我站在巨人的肩膀上。在研究工作中,任何一个项目从选题、试验研究或设计,到成果鉴定或申请专利,都要充分利用检索工具,占有详细资料,沟通情报信息,了解哪些工作前人已经做了,哪些目前正在做,什么人在做,进展情况如何。这样才能避免重复,少走弯路,保证工作在高水平上进行,并缩短周期,获得预期经济效益。在这方面国内外有许多成功的范例。

我国著名的青岛海尔集团建立了制冷技术、钢板、压缩机、发泡材料等技术信息库,据此进行大规模的新产品开发,一跃成为世界家电名牌。日本研制的高能加速器,由于充分利用了国外情报信息,研制费比国外同类装置节省 1/2,人力节省 3/4。山东南壁矿将引进涂料生产线改为专利检索后自行研制,节约经费近百万元。我国葛洲坝工程二江电站外线方案,由于及时搜集分析了大量国内文献资料,采用了高压架空线路方案,节约投资数百万元。

相反,如果继承和借鉴工作做得不好,闭门造车,就难免造成科研项目的重复以及人力、财力、物力和时间的浪费。例如在美国的“阿波罗登月计划中”,曾花费数百万美元,以研究解决甲醇对钛燃料箱的腐蚀问题。其实解决这个问题的方法在文献中早有记载:

只要在甲醇中加 2% 的水就行了。而检索到这一信息仅须 15 分钟,花费几美元。我国某市几位科研人员长期废寝忘食,研制成功一架熔点仪,满怀希望准备申报国家发明奖,但是在检索专利文献时发现几年前别人就已完成了同样的工作,并获得了专利权。又如某研究所费时 10 年研制成功一种以镁代银电镀新工艺,后来发现,美国某公司 1929 年就已申请了这方面的专利。现在该工艺已经超过了专利的保护期,任何人都可以自由使用,而且该项专利说明书就收藏在附近的研究所。据调查,我国每年各类科研、革新、改革中 56% 的课题重复国外已有的成果,而国内彼此重复的课题的约占 2/3。由于每年不必要的重复而给国家造成的损失约在 15 亿元以上。

五、促进信息资源的开发和利用

历代流传下来并且迅速涌现的信息资源,是一个巨大的财富宝库,是一种同能源、材料和人才一样重要的资源,直接影响到一个国家的科学、教育、文化和经济的发展。在当今的知识经济时代,国际间的竞争就是信息的竞争。信息业的发达程度,是衡量一个国家竞争力的重要标志之一。发展中国家由于信息意识薄弱,技术落后,不能及时、便利地获取全面而准确的信息;社会和个人的知识不能及时得到补充和更新,因而造成人才匮乏,经济滞后。从信息资源开发利用的角度来说,信息检索是人们打开知识宝库的一把金钥匙,是开发信息资源的有力工具。它能帮助人们传播和利用信息,把信息转化为社会物质财富或创造出更多的精神财富。

由于信息检索的上述诸多作用,我国也开始重视这项工作,并投入大量资金,引进新技术、新设备去武装它,建设有中国特色的信息高速公路,以支持科研和其他开发活动。

第二节 信息检索的研究对象和类型

一、信息检索的研究对象

信息检索作为一门科学或研究领域,是情报学、文献学、图书馆学、目录学等交叉产生的新学科,并且越来越多地从计算科学、数学、语言学等学科引进一些科学方法和科学手段。它的研究对象有下列四个方面:

1. 检索对象,即信息。包括信息的概念、性质、特点、类别、用途、收集方法及专门的查找工具。

2. 检索工具和检索系统,包括对其整体的全面考察和对某种工具的研究。

3. 检索语言和检索方法,这是信息检索的核心问题。

4. 检索策略、检索技术、服务方式及检索效果评价。

研究上述四个方面,目的是为了总结继承前人管理信息和利用信息的宝贵经验,建立和不断完善信息检索的理论,发展新方法、新的技术和新的服务领域,指导检索实践,使庞大的科技文献得到有效的控制和充分利用。

二、信息检索的类型

1. 按检索的方式分

(1)手工检索:手工检索是指利用各种检索工具,用人工翻检的方式查找到所需信息的方法。这种方式至今仍然是检索的基础,但费时费力且容易漏检,一般已很少使用了。

(2)计算机检索:就是利用计算机,根据某一种信息提问,在一定时间内,从经过加工处理并已存贮在计算机内的数据库中,查出

所需的信息。计算机检索目前也有多种检索方式,较常用的有微光盘检索和联机检索。现在国际互连网络(Internet)发展迅速,引发了一场信息检索革命,引导信息检索向一个更加广阔的领域突飞猛进。

2. 按检索的对象和结果分

按照检索的对象和结果分,信息检索有文献检索、数据检索、事实检索、全文检索和超文本检索等五种类型。

(1)文献检索:是指迅速、及时、准确、完整、经济、无重大遗漏地从浩如烟海的文献群中查出与特定的研究课题相关的文献线索或文献的过程。这种检索一般要利用书目、索引、文摘等检索工具,其检索结果为题录、文摘或全文及出处等线索性资料,属于一种相关性检索,一般不直接回答用户所提出的技术问题,只提供相关的文献供参考。

(2)数据检索:是将观察或实验得到的数据经过筛选、分析、整理和鉴定,以一定的顺序存贮在某种载体上,然后采用适当的方法或手段从中找出符合用户所需的数据的过程。

(3)事实检索:是将基本事实预先置入某种载体,当需要这种信息时再提取出来的过程。数据和事实检索都是确定性的检索,要向用户提供确切的数据和准确的事实。使用的检索工具主要是字典、词典、百科全书、年鉴、手册、类书、政书、名录等参考工具书。

(4)全文检索:是将文献全文存贮在检索系统中,当需要时,检索出各种信息单元。例如提供文献中的某个词或几个词所在的出处,出现的频率及所在句、段的上下文。全文检索是计算机检索的发展方向。

(5)超文本检索:一般要在国际互连网络上进行,通过遍布世界的4000多台服务器,将用户所需的各种文本、活动或静止图像、声音等信息进行屏幕显示或用某种载体下载。

不同的检索要求属于特定的检索类型。例如,查找电机优化设

计的参考资料即属于文献检索；查找变速器的设计参数就是数据检索；查找美国有哪些机构在进行超导研究即为事实检索；查找在《红楼梦》中，林黛玉哭过多少次、邓小平关于“一国两制”有哪些论述，则属于全文检索；在国际互联网上查找世界对印巴核试验的反应则可看作是超文本检索。

综上所述，广义的信息检索是指将信息按一定的方式组织和存贮起来，并根据信息用户的特定要求找出信息的过程和技术。狭义的信息检索仅指从信息集中查出所需要信息的过程。

三、信息、知识、情报、文献的概念

开设信息检索课的目的是为了增强学生的信息意识，学习信息检索的方法，提高开发和利用信息资源的能力，掌握读书治学的门径。由于历史的和现实的原因，在信息检索的过程中，信息、知识、情报、文献的基本概念及其相互关系在不同的教科书中有不同的解释，应该客观地加以界定。

1. 信息

对于信息的定义，众说纷纭，莫衷一是。近年来对信息较为一致的定义是：信息是世界上一切事物的运动状态、特征及其反应。它与事物同在，存在于整个自然界与人类社会。信息是人们认识世界，改造世界，取之不尽，用之不竭的宝贵资源。在当今知识经济主导的信息社会的时代，信息作为与自然资源并重的战略性资源，已经成为经济发展和社会进步的关键变量和重要支柱。

2. 知识

知识是人类对各种大量信息进行思维分析，加工提炼，并加以系统和深化而形成的结果。简而言之，理性化、优化和系统化了的信息即为知识。人类生活环境中普遍存在的信息，是构成知识的原料，这些原料经过人脑接收、选择、处理，最终组合成知识。新知识首先发生并存在于人的大脑中，这就是主观知识。如将大脑中的认