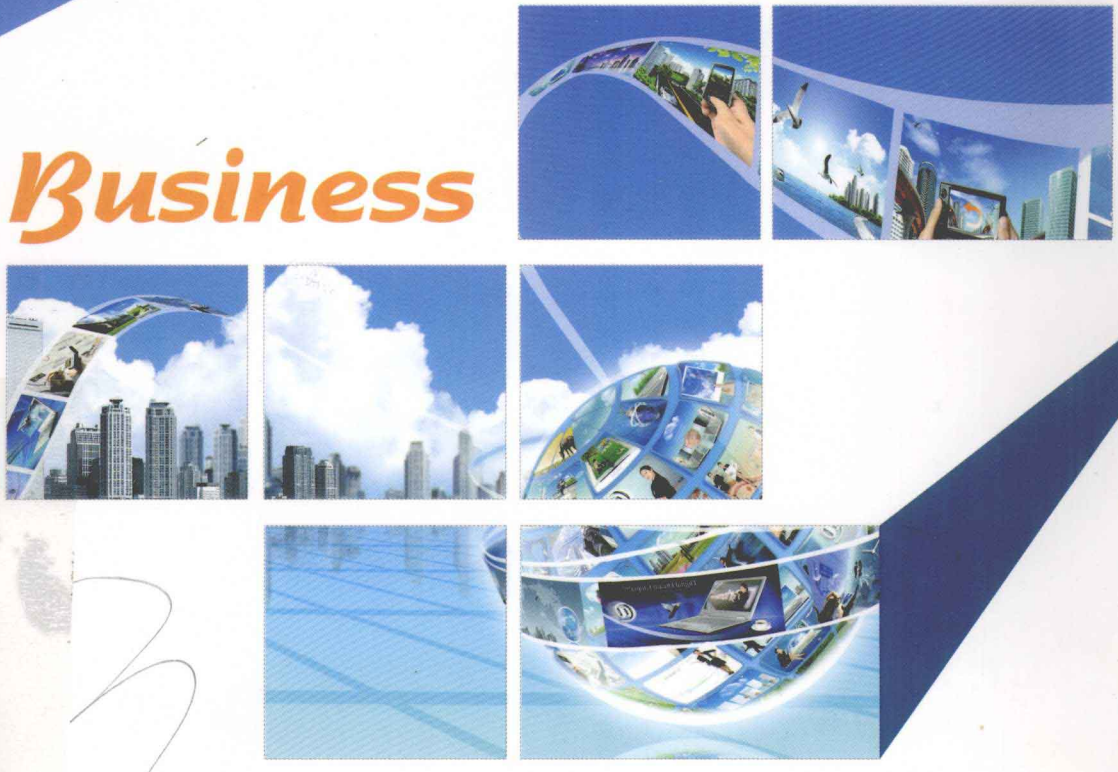
高等学校工商管理专业应用型本科系列教材

信息检索
——理论与创新

许福运　张承华　主编



高等教育出版社
HIGHER EDUCATION PRESS

高等学校工商管理专业应用型本科系列教

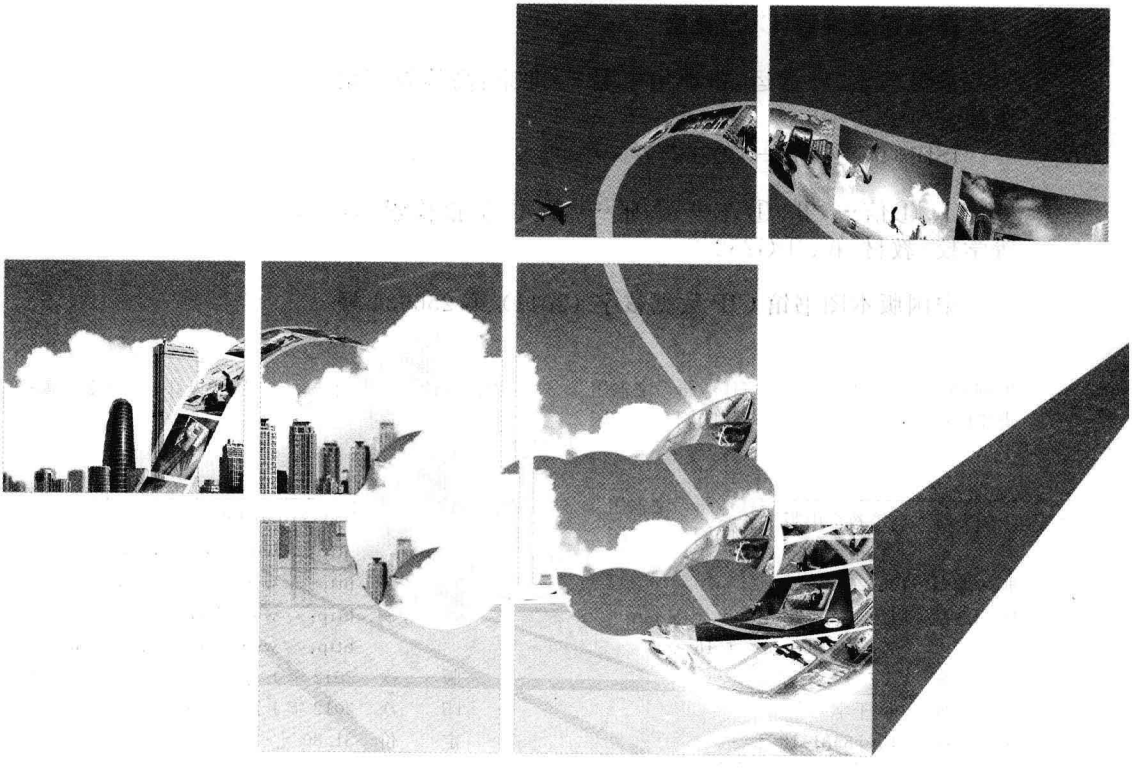
信息检索

Xinxi Jiansuo

——理论与创新

Lilun Yu Chuangxin

许福运　张承华　主编
马静玉　黄　睿　姜仁珍　副主编



高等教育出版社·北京
HIGHER EDUCATION PRESS　BEIJING

内容提要

快捷准确、及时有效地检索和利用信息，是网络环境下对信息检索提出的新要求，也是知识经济时代的劳动者必须具备的基本信息素养。本书为适应信息检索课程发展的需要，系统地阐述了信息检索的基本理论和方法，并适时地把信息检索领域的最新知识和成果充实进来。全书共分9章，包括：数字信息检索的基本理论和知识；中文数据库；国外几种在编排结构上具有代表性的著名综合性检索系统；经济管理专业重要检索系统；经济信息检索及专利标准信息检索；互联网信息检索以及搜索引擎的相关理论和知识。本书还对检索与创新一体化进行了阐述，对考研、就业、留学等实用信息检索进行了介绍，并对信息综合利用给出了典型案例。本书由多年从事高等学校"信息检索"课程教学和科研的教师编写而成，可作为广大文献信息工作者、教学科研管理和技术工作者用书，也可作为各类应用型本科高等学校相关专业开设"信息检索"课程的教材或教学参考书，对热衷于信息检索研究和学习的广大读者也十分适用。

图书在版编目(CIP)数据

信息检索/许福运，张承华主编. —北京：高等教育出版社，2012.1

ISBN 978-7-04-030762-7

Ⅰ.①信… Ⅱ.①许… ②张… 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（2011）第256934号

策划编辑 宋志伟　　责任编辑 宋志伟　　封面设计 张 楠　　版式设计 马敬茹
责任校对 刘春萍　　责任印制 刘思涵

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	唐山市润丰印务有限公司	网上订购	http://www.landraco.com
开　　本	787mm×960mm　1/16		http://www.landraco.com.cn
印　　张	21.75	版　　次	2012年1月第1版
字　　数	390千字	印　　次	2012年1月第1次印刷
购书热线	010-58581118	定　　价	31.80元

物 料 号　30762-00

前　言

21世纪的中国以创新为灵魂的知识经济正在成为社会发展的主流。在知识经济时代，信息广泛渗透到经济、科技、文化的各个学科领域乃至人类生活的各个方面，它是一种极其重要的经济资源、战略资源，是不可替代的生产要素和不断增值的社会财富。信息的获取和利用成为人类赖以生存与发展的一种本能。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，必须首先广泛地获取文献信息，在前人已经取得成就的基础上去进行新的探索。俄罗斯信息学家布留索夫认为，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海洋中找到知识的本领。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。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。掌握信息检索技术成为每个大学生和科研人员必备的基本技能之一。如何快捷、准确、及时、有效、经济地获取与自身需求相关和有用的信息，是知识经济和网络时代对信息检索提出的新要求，也是当代大学生必须具备的基本信息素质。因此，信息素质的培养日益成为世界各国教育界乃至社会各界所关注的理论与实践的重大课题。

近年来，世界信息环境发生着巨大变化。伴随网络化、数字化而产生的网络信息资源具有数量庞大、类型繁多、分布广泛的特点。数字资源、计算机网络存储和传递技术以及个性化的信息需求构成了新的信息环境。信息资源的组织、检索与利用模式正在发生根本性的变革。目前，培养学生在网络环境下获取和利用信息的能力已成为高等教育教学活动的基本要求之一，信息检索课程教学成为实现大学生获取与利用信息能力培养的重要教育环节。许福运教授组织全国信息检索课程教师及部分专家率先开展信息检索与创新教学研究，编撰出版了这部《信息检索》教材。通过现行的信息检索技能教育课程进行全方位的改革和探讨，形成一体化的教学模式，以拓宽信息检索教育的专业领域，推进并完善我国信息检索教育模式，实现大学生信息素养与创新能力的提高。

本书以提高大学生的信息素质和信息获取能力为出发点，突出信息检索的应用性，将侧重点放在叙述检索数字信息的方法和技巧上，向读者提供一部覆盖面广且较精炼实用的教材。

应用性是该书的主要特点。本书列举了很多检索实例并配有大量检索图示。凡涉及数字信息检索技术与方法的演示图片，都进行了反复实践操作验证，以确保所配实例操作过程的准确性。除实用性之外，新颖性也是本书所追求的目标。为此，本书的编著者们密切关注数字信息及信息检索工具的发展动态和最新成果，参考了许多最新资料，吸纳了一些新内容，尽可能反映最新的信息检索理论研究成果，以满足新世纪信息检索课程教学的需要。本书专门编写了检索与创新的章节，对创新思维和技巧进行了介绍并且给出了案例。这对于开拓读者的视野大有裨益，也是对检索与创新教育一体化的有益尝试。这是已出版的其他同类型教材中所不曾见到的。

本书共有九章，其中第一章由刘二稳编写，第二章由贺伟、姚伟编写，第三章由刘鹏编写，第四章由刘一农、姜仁珍编写，第五章、第六章由张承华、杨冰、黄睿编写，第七章由史建华编写，第八章由许福运、贺长伟编写，第九章由张承华、马静玉编写，全书由许福运、张承华统稿，许福运总纂。

该书的面世得到了高等教育出版社有关编辑的大力支持与帮助，在此表示衷心的感谢。在创作本书的过程中，听取了许多专家提出的宝贵意见，参考了大量文献资料包括部分网络资料，不便于一一标出，在此深表谢意。数字信息检索的理论和实践发展很快，新的理论和方法层出不穷。本书的编写，由于时间仓促，而编著者学识水平有限，因此疏漏、不足甚至错误之处在所难免，尚望广大读者不吝赐教，至为感谢。

编 者

2011 年 12 月

目　录

第一章　数字信息检索基础

第一节　数字信息资源概述

一、数字信息资源的概念

数字信息资源，狭义讲，亦可称为电子资源，指一切以数字形式生产和传递的信息资源。所谓数字形式，是以能被计算机识别的、不同序列的“0”和“1”构成的形式。数字资源中的信息，包括文字、图片、声音、动态图像等，都是以数字代码方式存储在磁带、磁盘、光盘等介质上，通过计算机输出设备和网络传送出去，最终显示在用户的计算机终端上。

随着互联网的发展，利用网络传递的数字信息资源的数量每年都以几何倍速增长，我们把这一类数字资源均称为网络信息资源（network information resources）。网络信息资源目前在数字信息资源中已经占有绝对比例。除此之外，到目前为止，仍然存在着大量仅在本地计算机上使用、没有通过网络传递的信息资源，如只用于单机的光盘或机读磁带数据库等，我们把这一类资源也归为数字信息资源。

二、数字信息资源的类型

数字信息资源的类型可以从数字信息的记录形式、时序、出版形式、制作形式四个方面进行划分。

（一）按记录信息的形式分

根据记录信息的形式，数字信息资源可划分为文本类、图形（图像）类和声（视）频多媒体类三种。

1. 文本类数字信息资源

它是用语言、文字记录的信息资源，是最重要的信息资源，即文献。它是使用计算机文字处理软件产生的文件保存格式。其主要格式有：①TXT 文本格式；②DOC 文件格式；③HTML 超文本格式；④PDF 格式。

虽然 DOC、HTML 和 PDF 格式都可以包含图形、声音等多媒体信息，但由于

它主要以文字、文本信息为主，图形、声音是辅助性的，因而我们仍然把它看成是文本数字信息。

2. 图形（图像）类数字信息资源

图形（图像）类数字信息资源主要包括各种照片、绘画、图谱、图片、图纸、图表等。目前常见的图形（图像）格式大致分为两大类：一类为位图；另一类为描绘类、矢量类或面向对象的图形（图像）。前者是以点阵形式描述图形（图像）的，后者是以数学方法描述的一种由几何元素组成的图形（图像）。一般来说，后者对图像的表达细致、真实，缩放后图像的分辨率不变，在专业级的图像处理中运用较多。

图形（图像）类数字信息资源的文件保存格式主要有：①BMP 格式；②JPG 格式；③GIF 格式；④TIFF 格式；⑤PSD 格式等。

3. 声（视）频多媒体类数字信息资源

多媒体数字信息是指既用文字、图、表、符号记录，也用声音、影像等记录，是集文字、声音、图形（图像）、影像于一体的数字信息资源。目前较为流行的声（视）频多媒体文件格式有：①WAVE，扩展名为 WAV；②MOD，扩展名为 MOD、ST3、XT 等；③MPEG-3 和 MPEG-4，扩展名分别为 MP3、MP4 等；④Real Audio，扩展名为 RA 等。

（二）按文献的时序形式及有序化程度分

1. 零次文献

零次文献也称零次信息，指未经正式发表或不宜公开和大范围交流的比较原始的素材、底稿、手稿、书信、工程图纸、考察记录、实验记录、调查稿、原始统计数字，以及各种口头交流的知识、经验、意见、论点等。

此类文献的形式为手抄本、油印件、复印件等；电子形式为内部录音、录像、E-mail、BBS 帖子、电子文档等。

2. 一次文献

一次文献即原始文献，指反映最原始思想、成果、过程以及对其进行分析、综合、总结的信息资源，如事实数据库、电子期刊、电子图书、发布一次文献的学术网站等。用户可以从一次文献中直接获取自己所需的原始信息。

此类文献的印刷形式主要包括图书、期刊和报纸、科学考察报告、研究报告、会议论文、学位论文、专利说明书、技术标准、政府出版物、产品样本等；电子形式包括实时数据库、电子期刊、电子图书、电子预印本和发布一次文献的正式学术网站等。

3. 二次文献

二次文献也称二次信息，习惯上又称检索工具，是根据实际需要，按照一定的科学方法，将特定范围内的分散的一次文献进行筛选、加工、整理，使之有序化而形成的文献。由于它能较为全面系统地反映某学科、某专业的文献线索，因而是

检索和评价一次文献的便捷工具。

此类文献的印刷形式有书目、文摘、题录、索引等；电子形式有二次文献数据库、搜索引擎等。其中二次文献数据库是在传统检索工具（如书目、文摘、题录、索引）基础上形成和发展起来的数据库。

4. 三次文献

三次文献也称三次信息，是指通过二次文献提供的线索，选用一次文献的内容，进行分析、综合、研究后而编成的文献。一般包括专题述评、专题调研、动态综述、进展报告、学科年度总结等。此类文献的印刷形式和电子形式基本重合，都包括综述、述评、字词典、百科全书、年鉴、标准、数据手册等。

（三）按文献的出版形式分

数字信息资源的出版形式是图书馆和信息服务机构存放、管理和提供服务的依据。根据文献出版形式的类型和特点不同，文献型数字信息资源分为以下几种类型。

1. 图书

图书是最早的文献类型之一，至今仍占据文献的主导地位。它具有内容成熟、知识系统完整，但传递知识信息相对较慢，编辑出版的周期较长等特点。图书根据功能不同分为阅读类和工具类。阅读类图书包括各种教科书、专著、文集等。工具类图书包括各种百科全书、年鉴、手册、词典、指南、名录、图册等。

图书著录的主要外部特征是：书名、著者、出版社名称、出版地、出版时间、总页数和国际标准书号（ISBN）。图书辨识的直接关键词是“出版（社、者）”，英文词是 press、publication（pub.）、publisher。例如：

Computer Simulation of Electrionic①, R.Raghuram②, New Delhi, India: Weiley③ (1989)④ 246pp⑤, [8122401112]⑥

注：①书名；②著者；③出版地、出版者；④出版日期；⑤图书总页数；⑥国际标准书号。

2. 期刊

期刊又称杂志（journal）、连续出版物（serials），指定期或不定期出版的有固定名称的连续出版物。它们有连续的卷期或年月顺序号，具有周期短、反映新成果及时、内容新、学术性强、信息量大等特点。

期刊的类型常用冠名：acta（学报）、journal（杂志）、chronicles（纪事）、annuals（年刊）、bulletin（通报）、transactions（汇刊）。

期刊文献著录的主要外部特征是：论文题名、著者、刊名、卷号（Vol.）、期号（No.）年月、起至页码、国际标准刊号（ISSN）。其中：卷号（Vol.）、期号（No.）年月、起至页码、国际标准刊号（ISSN）是辨识期刊文献的主要外部特征。上述期刊类型

的常用冠名也是辨识期刊的直接关键词。例如：

J.Pressure Vessel Technol. Trans, ASME① V112 n.4② Nov 1990③ p410-416④

注：①刊名缩写；②卷期；③出版年月；④起止页码。

3. 会议文献

会议文献是指在学术会议上宣读或交流的书面论文。会议文献一般有会前、会中和会后文献三种类型。其中会后文献是主要的文献类型，通常以期刊、论文集、会议记录等形式出版。会议文献的特点是：文献论题集中，内容新颖，学术性强，能反映一个国家、一个地区或某一科学技术领域的最新成就、最高水平和发展趋势。许多最新的研究成果往往首先在会议上发表，所以会议文献成为了解各国科技发展水平和发展动向的重要文献源，因而受到科学界的高度重视。

会议和会议文献常用的主要名称有大会(conference)、小型会议(meeting)、讨论会(symposium)、研讨会(seminar)、会议录(proceeding)、单篇论文(paper)、汇报(transaction)等。

会议文献著录的主要外部特征是：论文题名、著者、编者、会议名称或会议论文集名称、会议地或主办国、会议时间、论文在会议论文集中起至页码、会议论文编号。其中：会议名称或会议论文集名称、会议地或主办国、会议时间、论文在会议论文集中起至页码、会议论文编号是辨识会议文献的主要外部特征。上述会议和会议文献常用的主要名称也是辨识的直接关键词。例如：

Proceedings Fourth Annual Symposium on Logic Computer Science①(Cat. No.89 CH2753-2)② Paific Grove, CA.USA, 5-8 June 1989③(Washington, D.C.USA: IEEE Computer, Soc Press 1989 p263-72)④

注：①会议及会议录名称；②订购号；③会议地点及时间；④会议录出版单位、地址及出版年份和页码。

4. 学位论文

学位论文是高等院校的学生为了获取一定的学位资格而撰写的学术性研究论文，如博士论文、硕士论文、学士论文等，其特点是具有学术性和独创性。大多数国家采用学士(Bachelor)、硕士(Master)和博士(Doctor)三级学位制。通常所讲的学位论文，主要指博士、硕士论文及优秀学士学位论文。

学位论文除在本单位被收藏外，通常也被本国的国家图书馆收藏。例如，在我国，国家科技文献中心(NSTL)、中国科技信息研究所、万方数据、CNKI(清华同方)都集中收藏和报道国内各学位授予单位的博士、硕士学位论文。

学位论文著录的主要外部特征是：学位名称、导师姓名、学位授予机构、学位授予时间等。学位论文辨识的直接关键词是“学位论文”和“学位名称”，英文词

是 doctoral dissertation 和 M.S、M.B.A、.Ph.D.、D.E、D.S. 等。例如：

Allen, B.[①], Learning Body Shape Models from Real-World Date[②], ph.D.Thesis[③], 2005[④]. (pdf)[⑤]

注：①作者；②论文题目；③论文级别；④时间；⑤原文格式。

5. 科技报告

科技报告是科技人员从事某一专题研究所取得的成果和进展的实际记录。科技报告一般都有编号，且单独成册。科技报告反映的是新兴科学和尖端科学的研究成果，内容新颖，专业性强，能代表一个国家的研究水平，各国都很重视。目前，美、英、德、日等国每年产生的科技报告达 20 万件左右，其中美国占 80%。美国政府的 AD、PB、NASA、DOE 四大报告在国际上最为著名。

(1) PB(Publishing Board)报告。由美国商务部国家技术情报服务处(NTIS)出版发行，报告的内容侧重于各种民用科学技术、生物医学。

(2) AD(ASTIA Document)报告。原来为美国武装部队技术情报局(Armed Services Technical Information Agency, ASTIA)出版的文献，即 ASTIA Document 报告。现今 AD 含义已变为入藏文献(Accessioned Documents)，主要收录军事科技方面的文献资料。

(3) NASA 报告。NASA 是美国航空航天局(National Aeronautics and Space Administration)的简称。内容除航空航天技术以外，涉及许多相关学科，在一定程度上成为综合型科技报告。

(4) DOE 报告。DOE 报告是美国能源部(US Department of Energy)出版的报告，收录能源部所属实验室、能源技术中心和情报中心以及合同单位发表的科技报告，内容涉及核能与其他能源，包括矿物燃料、太阳能以及节能、环境和安全等内容。

科技报告具有保密的特点，因而不易获取。在我国，国家图书馆、国防科技信息研究所和上海图书馆的科技报告相对比较完整。

科技报告文献著录的主要外部特征是：报告名称、报告号、研究机构、完成时间等。例如：

Report KFKI—1983—570[①], Hungarian Acad, Sci., Budapest[②] (1983)[③], 15pp[④]

注：①科技报告编号；②收集或编写科技报告机构地址；③公布时间；④报告页数。科技报告有“Report”这个特征词。

6. 专利文献

专利文献指根据专利法公开的有关发明的文献，主要为专利说明书，也包括

专利法律文件和专利检索工具。专利文献具有新颖性、创造性和实用性的特点，且范围广泛，出版速度快，格式规范，有助于科技人员借鉴国际先进技术，避免重复劳动。专利文献是一种实用性很强的技术资料。

专利文献的主要外部特征有：申请号、公开号、申请人（专利权人）、发明（设计）人、申请日、公开（公告）日等。例如：

……Patent no：US 4202737……

注：US 是美国的专利号代码，专利文献有“patent”这个特征词。

7. 标准文献

标准文献是对工农业新产品和工程建设的质量、规格、参数及检验方法所做的技术规定，是人们在设计、生产和检验过程中共同遵守的技术依据。它是一种规章性的技术文件，具有一定的法律约束力。按批准机构级别和适用等级可分为国际标准、国家标准、部颁标准（行业标准）和企业标准四个等级。标准的主要收藏单位是省级以上的技术监督研究所和科技信息所。

标准文献都有标准号，通常由国别（组织）代码 + 顺序号 + 年代组成。我国的国家标准分为强制性的国标（GB）和推荐性的国标（GB/T）；行业标准代码以主管部门名称的汉语拼音声母表示，如 JT 表示交通行业标准；企业标准编号为：Q/ 省、市简称 + 企业名代码 + 年份。

标准文献著录的主要外部特征是：标准级别、标准名称、标准号、审批机构、颁布时间、实施时间等。标准文献辨识的直接关键词是“标准”（standard）与“标准号”。例如：

GB① 50411② -2007③

注：①标准代号（中国国家标准）；②顺序号；③年份。

8. 政府出版物

政府出版物指各国政府部门及其所属机构颁布的文件，如政府公报、会议文件和记录、法令汇编等。所包含内容范围广泛，几乎涉及整个知识领域，但重点在政治、经济、法律、军事等方面。政府出版物按其性质可分为行政性文献和科技性文献，它具有正式性和权威性的特点。

西方国家多设有政府出版物的专门出版机构，如英国的皇家出版局（HMSO）、美国政府出版局（GPO）等。中国的政府出版物大部分是由政府部门编辑，由指定出版社出版。

9. 产品资料

产品资料又称产品说明书，是对一种产品的性能、规格、构造、用途、使用方法

等所作的说明。产品资料技术比较成熟、数据可靠，常附有外观照片与结构图，直观性强。但产品资料的时间性强，使用寿命较短。产品资料是技术人员设计、制造新产品的一种有价值的参考资料。

10. 技术档案

技术档案指生产建设、科技部门和企事业单位针对具体的工程或项目形成的技术文件、设计图样、图表、照片，原始记录的原本及复印件。包括任务书、协议书、技术经济指标和审批文件、研究计划、研究方案、试验记录等。它是生产领域、科学实践中用以积累经验、吸取教训和提高质量的重要文献。科技档案具有保密性，常常限定使用范围。

第二节　数字信息检索理论

一、数字信息检索的概念和原理

（一）数字信息检索的概念

数字信息检索是指人们在计算机或网络终端上，使用特定的检索指令、检索词和检索策略，从计算机检索系统的数据库中检索出所需要的信息，再由终端设备显示、下载、拷贝、打印的过程。广义的数字信息检索包含信息存储和信息检索两个部分（见图 1-1）。

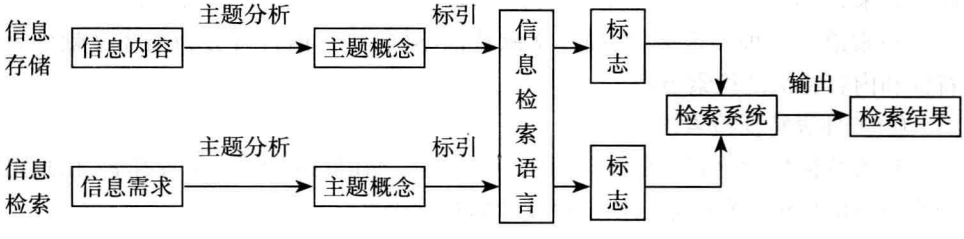


图 1-1　数字信息检索的全过程

（1）信息存储，就是指标引人员对文献内容进行主题分析，即把文献包含的信息内容分析成若干能代表文献主题的概念，并用词表、分类表等规范标志的检索语言对文献主题进行标引，同时把入选文献中的其他特征标志（标题、著者、文摘、原文出处等）按所选数据库结构的索引结构一起输入到计算机文献检索系统中进行存储，编制成一系列索引文档数据库和文摘信息数据库。

(2) 数字信息检索是检索者对检索课题进行主题分析,明确检索范围,形成能代表信息需求的若干主题概念,把这些主题概念转换成计算机信息检索语言,即用数据库、检索工具书对各概念选词并进行概念逻辑组配,编制成文献检索提问式,再在所选的合适数据库中用检索系统规定的指令输入到计算机,通过检索软件在数据库中运行,将信息需求的主题概念和数据库内文献主题标志进行匹配,找到命中文献。

(二) 数字信息检索的原理

无论是数字信息还是传统信息,其检索原理基本相同,即对信息集合与需求集合的匹配与选择,也就是检索提问标志与存储在数据库中的文献标引标志进行比较,两者一致或信息标引的标志包含着检索提问标志,则具该标志的信息就从数据库中输出,输出的信息就是检索命中的信息(见图 1-2)。

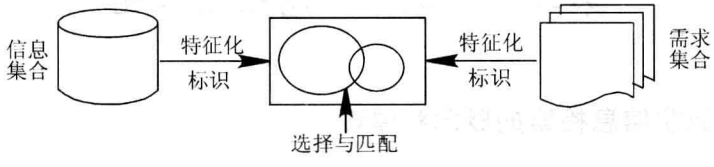


图 1-2 数字信息的检索原理

二、数字信息的检索语言

检索语言就是信息组织、存储与信息检索时所用的语言,它是把存储与检索联系起来,使信息处理人员和检索人员共同遵守的语言(一种人工语言)。

检索语言的种类很多,按描述文献特征不同,检索语言可分为描述文献外表特征和内容特征的检索语言。

(一) 外表特征检索语言

外表特征检索语言是依据信息的外表特征,如信息的题名、作者姓名、信息出处等作为信息标引和检索的依据而设计的检索语言。

1. 题名检索语言

这是指以书名、刊名、篇名、论文题名为标志的检索语言。题名检索语言一般规定:题名索引按字顺排列,如西文题名中的虚词不作索引,实词按字母顺序排列,中文按汉语拼音字母顺序或汉字的笔画笔形排列。

2. 著者检索语言

这是指以作者、译者、编者等信息的责任者的姓名或团体组织名称为标志的检索语言。这种语言一般要求:著者的姓名用姓在前、名在后的形式,而且姓要用

全称，名要用缩写。但各个数据库的要求则不尽相同。因此，检索时要参考检索工具的使用说明。

3. 序号检索语言

这是指以信息特有的序号为标志的检索语言，如专利号、技术标准号、化学文摘号等。使用这种检索语言，也要注意各检索工具的差别。

（二）内容特征检索语言

1. 分类检索语言

分类检索语言是指按照学科范畴及知识之间的关系列出类目，并用数字、字母符号对类目进行标志的一种语言体系，也称分类法。目前常用的分类法有《中国图书馆图书分类法》（简称《中图法》）、《美国国会图书馆分类法》、《杜威分类法》、《国际专利分类法》等。例如《中图法》将所有的知识分为22个大类，并且用不同的字母标志不同的学科，构成一个知识体系，如：

……	O 数理科学和化学
F 经济	P 天文学、地球科学
G 文化、科学、教育、体育	Q 生物科学
H 语言、文字	R 医药、卫生
I 文学	S 农业科学
……	……

其中每一个大类又可以细分成若干个二级类目，二级类目还可以再细分。例如经济又可以划分为：

F0 经济学	F3 农业经济
F1 世界各国经济概况、经济史、经济地理	F4 工业经济
F2 经济计划与管理	……

“工业经济”又可以进一步划分为“工业经济理论”、“世界工业经济”、“中国工业经济、”“各国工业经济”等。这些类目还可以再层层划分，每一级类目都用“字母＋数字”形式进行标志。

2. 主题检索语言

由主题词构成，是直接以代表文献内容的主题概念作为检索标志，并按其字顺组织起来的一种检索语言。根据词语的选词原则、组配方式、规范方法，主题语言可分为标题词语言、单元词语言、叙词语言和关键词语言。

（1）标题词，是一种先组式的规范词语言，即在检索之前已经将概念之间的关系组配好。具有较好的通用性、直接性和专指性，但灵活性较差。

（2）单元词，是一种最基本的、不能再分的单位词语，亦称元词，从文献内容中抽出，再经规范，能表达一个独立的概念。例如“信息检索”是一个词组，“信息”