



高等院校“十三五”应用型规划教材

XINXI JIANSUO
SHIJIAN JIAOCHENG

信息检索实践教程

主 编 于喜展 孙志梅



南京大学出版社





高等院校“十三五”应用型规划教材

信息检索实践教程

主 编 于喜展 孙志梅

微信扫一扫



教师服务入口



学生服务入口



南京大学出版社

图书在版编目(CIP)数据

信息检索实践教程 / 于喜展, 孙志梅主编. — 南京:
南京大学出版社, 2017. 8

高等院校“十三五”应用型规划教材

ISBN 978-7-305-19084-1

I. ①信… II. ①于… ②孙… III. ①情报检索—高等学校—教材 IV. ①G254.9

中国版本图书馆 CIP 数据核字(2017)第 175573 号

出版发行 南京大学出版社

社 址 南京市汉口路 22 号 邮 编 210093

出 版 人 金鑫荣

丛 书 名 高等院校“十三五”应用型规划教材

书 名 信息检索实践教程

作 者 于喜展 孙志梅

责任编辑 张亚男 武 坦 编辑热线 025-83597482

照 排 南京南琳图文制作有限公司

印 刷 常州市武进第三印刷有限公司

开 本 787×1092 1/16 印张 12.25 字数 306 千

版 次 2017 年 8 月第 1 版 2017 年 8 月第 1 次印刷

ISBN 978-7-305-19084-1

定 价 29.00 元

网址: <http://www.njupco.com>

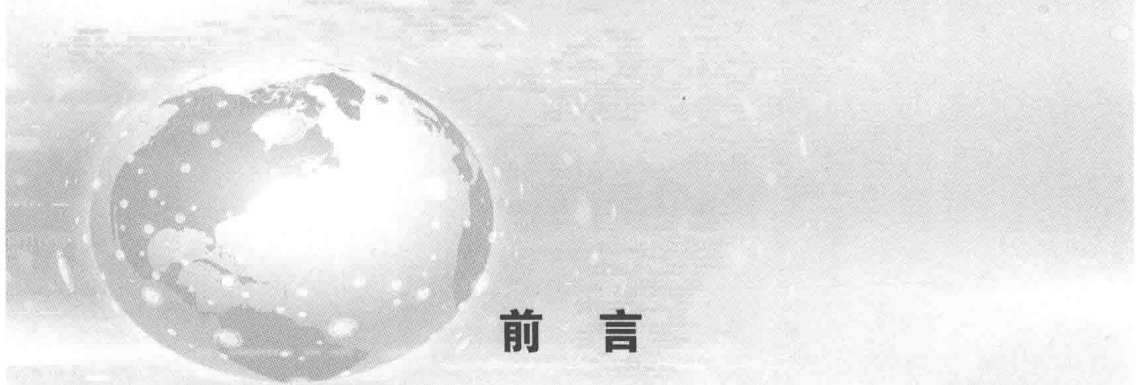
官方微博: <http://weibo.com/njupco>

微信服务号: njuyuexue

销售咨询热线: (025) 83594756

* 版权所有, 侵权必究

* 凡购买南大版图书, 如有印装质量问题, 请与所购
图书销售部门联系调换



前言

信息检索课是培养学生掌握文献信息的检索与利用,不断提高学生信息素质和能力的一门科学方法课。自1984年教育部印发《关于在高等学校开设“文献检索与利用”课程意见》的通知后,信息检索课被正式列入我国高校文理工科各专业的教学计划,也成为高校图书馆开设的唯一课程。随着社会信息化程度的不断提高,信息检索教育成为大学生信息素质培养和提高的一个重要途径。因而,许多高校已将其列为学生的必修课或选修课。通过该课程的教学,培养学生获取信息、开拓思维、勇于创新的才能,提高学生分析问题和解决问题的能力。

本教材主要为满足信息检索课程的教学、培养提高学生的信息素质水平和实践技能的需要而编写的。教材遵循普通高等院校对人才培养目标的要求和课程特点,以大学图书馆资源为基础平台,兼顾开放网络资源,按照“理论够用、注重实践操作”的原则,由一般到具体,虚实并重,理论与案例结合,力求:①理论结合实际,博采众长,充分吸收国内外相关实践经验和最新研究成果;由浅入深、循序渐进、简明扼要,便于初学者理解掌握。②在内容取舍与安排上,每章要包含基础理论知识、检索方法与技巧、检索应用案例等内容,做到体系完整而又重点突出,便于学习者在掌握学科体系的同时把握住重点和难点。③注重实践能力的培养,各章通过检索案例分析(具体到操作步骤)让学生掌握检索方法和技巧,每章后附有实践技能训练方案,以强化实际操作,提高实践能力。

本教材在编写过程中得到了山东科技大学图书馆长李家祥教授和副馆长张福俊副研究员的大力支持和帮助,在初步策划、内容选取、出版使用等方面给予联系和把关。具体由于喜展、孙志梅担任主编,完成全书的统稿审核;李默、李瑞芹、王亚博、徐静(齐鲁师范学院)担任副主编。编者具体分工如下:第一章(李默、于喜展);第二章(于喜展);第三章(于喜展、李瑞芹);第四章(王亚博);第五章(孙志梅);第六章(李瑞芹);第七章(孙志梅、徐静)。

编者在写作过程中,直接或间接地参阅和借鉴了国内外大量的有关专著、教科书和论文,因数量很多,难以一一列举,在此一并致谢。

由于编者水平所限,书中缺点、疏漏和错误在所难免,恳请广大读者批评指正。

编者

2017年6月



目 录

第一章 信息资源与信息素养	1
第一节 信息资源概述	1
第二节 主要信息资源	6
第三节 信息素质与创新能力	12
第四节 综合检索实例及分析	16
第二章 图书馆资源与利用	24
第一节 图书馆的职能与服务	24
第二节 馆藏资源检索方法	25
第三节 电子图书资源的获取	27
第四节 文献传递与馆际互借	31
第五节 综合检索实例及分析	34
第三章 信息检索的基本原理	47
第一节 信息检索基础知识	47
第二节 信息检索策略与步骤	54
第三节 综合检索实例及分析	62
第四章 常用数据库资源检索	75
第一节 中国学术资源总库(CNKI)	75
第二节 Ei village 数据库	82
第三节 Web of Science 数据库	86
第四节 ScienceDirect 数据库	97
第五节 综合检索实例及分析	99

第五章 特种文献检索	110
第一节 特种文献检索概述	110
第二节 专利文献检索	111
第三节 标准文献及检索	132
第四节 综合检索实例及分析	140
第六章 网络信息资源检索	148
第一节 网络信息资源基础知识	148
第二节 搜索引擎及检索技巧	150
第三节 开放学术信息资源及其利用	159
第四节 综合检索实例及分析	162
第七章 学术论文写作规范	169
第一节 学术论文写作基础知识	169
第二节 学术论文的写作与发表	172
第三节 毕业论文的写作规范	182
关于印发《发表学术论文“五不准”》的通知	186
发表学术论文“五不准”	187
参考文献	189



第一章 信息资源与信息素养

第一节 信息资源概述

资源是一种能够创造物质财富的东西;信息是物质存在的一种方式、形态或运动状态,也是事物的一种普遍属性;信息资源是人类认识和改造客观世界的知识源泉。在现代社会中,物质资源、能量资源与信息资源共同构成一切事物的三大要素,是现代社会发展的三大支柱,它们互相结合、互相依存。其中,物质资源是人类的物质基础,能量资源给予人类动力,而信息资源则是人类智慧和知识的结晶。

一、信息与信息资源

(一) 信息的定义与属性

1. 信息的定义

信息科学的重要学者维纳在《控制论——动物和机器中的通信与控制问题》中提出了关于信息的两个经典定义:其一,“信息既不是物质,也不是能量,信息就是信息,不懂得它,就不懂得唯物主义”;其二,“信息是用于适应外部世界,并且在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名称”。

2. 信息的属性

信息是宇宙万物的时空序,这里的序包括有序和无序,信息科学就是研究信息的属性及其与物质和能量之间的依存和转换关系的学科。信息有多种属性,其中有四个属性是最基本的,它是信息不同于物质和能量的根本区别。

1) 反映性

宇宙的任何实体都是信息的本原,世界总是在其不同时空点下形成具有实体结构状态的物质—能量—信息关系。信息来源于客观世界的物质实体,以及实体之间、实体和反映对象之间的关系,但绝对不是物质实体本身,因为没有反映性就没有信息,反映性是信息的最基本属性。信息不但反映实体的属性、状态、特征、结构和动态过程,而且反映它们之间的依存和转换关系。只有将信息首先理解为本原的反映,才能解释为什么信息总是与特定的对象或环境相关,而没有这种相关就没有信息。

2) 无量纲

信息没有独立的空间维度,实体对象发送信息,空间属于实体,物理实体多大或多小,信

息只是一种反映和表现;信息也没有独立的时间维度,实体对象发送信息,时间也属于实体,虽然一个对象的信息模型是与时间无关的,但如果必需,时间又可以与对象关联。正是由于信息具有无量纲的属性,我们才可以对下列现象进行解释:一个硬盘存满了 0、1 或它们的任意组合,与一个格式化了了的空白硬盘相比,在重量上完全相同。

3) 不守恒

根据守恒的自然定律,物质和能量既不能再生,也不能消灭,它只能进行转换。而信息世界是不守恒的,因为信息可以再生或消灭。比如,语言信息、科学知识、病毒信息等等都遵循不守恒的自然定律,它不但可以增加、减少或转换,而且可以消灭或再生。只要承认信息可以增加或减少、消灭或再生,信息就是不守恒的。事实上,只要承认信息的反映性和无量纲属性,就必然要承认信息的不守恒属性。

4) 可共享

信息是对实体对象的反映、显示或映射,信息对实体对象的依赖是绝对的。但信息对载体的关系可以相对独立,同一信息可以负载于不同类型的载体,同一类型的载体可以负载不同的信息。信息的可共享来源于信息对载体的相对独立的关系。从信息的相对独立性上看,信息被多个信宿共享和重复使用,对信息本身而言没有任何损失。信息是可共享和可重复使用的,而实体则是独占性的。比如有两个人,每人各有一个苹果(实体),他们交换后,每人还是一个苹果;但若每人都各有一条信息,经过交换后,每人就拥有两条信息。人类正是依靠信息与载体的相对独立性,尤其是信息的可共享性,才得到不断的进化和发展。

(二) 知识、数据、事实和文献

从狭义上讲,知识是人们通过学习、发现或感悟到的对世界的认识的总和,是人类认识的结晶。从广义上讲,知识是一种数据、事实和信息的有组织的认识、价值观、洞察力的体系。利用知识所构成的框架,人们可进而评价和吸收新的经验与信息。因此,知识是建立在数据、事实和信息基础之上的。

数据泛指对客观事物的数量、属性、位置及其相互关系的抽象表示,以适合于用人工或自然的方式进行保存、传递和处理。传统的数据管理是数据项、记录和文件的管理,现代的数据管理是数据库管理。随着信息技术的不断发展,计算机信息处理范围的扩大,数据的含义也扩大了,存储的数据也在向着“海量”方向发展。

事实是客观事物的存在状态。事实的最大特点是客观性,即客观、真实。事实在一般场合下是一种未经处理过的客观事物状态,但它也可能被人为改变。即使被人更改或变换,它仍然只属于客观的不以人的认识、意志为转移的事物状态。

《文献情报术语国际标准》将文献定义为:“在存储、检索、利用或传递信息的过程中,可作为一个单元处理的,在载体内、载体上或依附载体内而存储有信息或数据的载体。”简而言之,文献是记录有信息、数据、事实和知识的一切载体。构成文献有三个最基本的要素:① 构成文献内核的信息、数据、事实和知识;② 负载信息、数据、事实和知识的物质载体;③ 记录信息、数据、事实和知识的符号系统。

信息、数据、事实、知识是文献的内容,离开了信息、数据、事实和知识便不能称之为文献。物质载体是符号赖以依存的寄主,符号系统是信息、数据、事实和知识的携带者,物质载体和符号系统是信息、数据、事实和知识进入文献记录的前提和具体的形式、方法和过程。

文献具有信息性、传递性和动态性特征,这些特征是社会信息学、信息计量学和信息检索研究的重要内容。

(三) 信息与相关概念的区别

根据卡尔·波普的“三个世界”的理论,信息可以分为三大类:第一类是有关客观物质世界的信息,即本体论意义上的本原信息,它反映事物运动的状态及其变化的方式;第二类是有关人类主观精神世界的信息,即认识论意义上的抽象信息,它反映人类能感受的事物运动状态及其变化方式,处于意识、思维状态;第三类是有关客观意义上概念世界的信息,即主体论意义上的累积信息,它反映人类所表述的事物运动状态及其变换方式,用语言、文字、图像、影视、数据等各种载体来表示,汇成一个实在的自主的“信息世界”。以“三个世界”的理论来研究信息、数据、事实、知识和文献,它们之间存在以下关系。

信息存在于全部的三个世界(主观世界、客观的物质世界、客观的概念世界)中,知识存在于主观世界和客观的概念世界,但不存在于客观的物质世界,因此知识包含于信息之中。数据、事实存在于主观世界、客观的物质世界,但不存在于客观的概念世界中,因此数据、事实包含于信息之中。文献存在于客观的概念世界,但不存在于主观世界、客观的物质世界,因此文献也包含于信息之中。在信息及其相关概念中,知识离信息最近,二者的区别主要有以下三点:

- (1) 组织性区别。信息是零散的、片断的、特殊的;知识是结构的、连贯的、普遍的。
- (2) 时间性区别。信息是适时的、短时的,甚至是短命的;知识是持久的、无限的。
- (3) 空间性区别。信息是跨越式的、流变性的;知识是定位式的、确定性的。

简言之,信息像是无数的细胞,知识则像是一个有机体;信息可以设想为各种瞬时变化,知识则反映某种变化的存在;信息可以看成是多种变换,知识则跟踪某种变换的过程。由于人类获取信息的根本目的是获得客观世界的知识,所以,信息与知识的关系可以简化为“信息—知识—信息—知识”这样一种循环转换关系之中。

(四) 信息资源的概念

信息资源是指人类在社会实践中,对信息获取、筛选、处理、传输并存储在一定的载体上进行利用,最后得到的可以为人类创造新的精神财富和物质财富的信息集合。狭义的信息资源指的是信息本身,而广义的信息资源还包括与信息紧密相连的信息设备、信息人员、信息网络、信息系统等。由于信息本身既不是物质,也不具有能量,因此,任何信息都离不开一定的物质载体和能量,信息总是以一定类型或形式的载体作为其存在和传递的物质基础。

信息载体又可分为信息表征载体和物质载体两类,其中,信息表征载体是指信息内容赖以存在的文字、符号、声波、光波等用以记录信息的载体,亦称第一载体;物质载体则是指信息载体赖以存在和传递的纸张、胶片、磁带、光盘等物质材料,亦称第二载体或硬载体。只有第一载体和第二载体有机结合,才能保证信息内容得到完整保存和有效传递。而随着信息科学、材料科学的不断发展,信息载体也在相应地发展变化。

二、信息资源的特征

(一) 信息资源具有内容丰富、数量庞大的特点

各种信息收藏机构收藏的信息资源包括传统信息资源(缩微胶片、各类光盘、纸质文献

等)和网络信息资源,这些信息资源内容十分丰富,而且数据量特别庞大。随着互联网的迅猛发展,网络已经成为全世界最大的信息资源产生基地,这些网络信息资源包括成千上万种报纸、期刊,以及个人、公司、学校以及政府的详细信息,信息具有跨地区、分布广、多语种、高度共享的特点,并且随着时间的推移,信息资源数量呈指数增长。

(二) 信息资源还具有类型繁多、范围增大、用途扩大的特点

信息资源建设速度越来越快,相应的信息类型也越来越多,其包括了印刷型出版物、电子期刊、图书、工具书及大量的联机数据库、书目数据库、软件资源等种种形式。信息资源呈现出多种表现形式,比如文本信息中就包含了文字、符号、表格等,另外还有图像、声音、视频和动画等信息。除了传统的信息传递方式外,人们现在还通过网络进行信息交流。

(三) 信息资源具有更新速度加快、半衰期缩短的特点

科学技术的快速发展,使得各类文献资料和互联网信息不断地淘汰和更新,新技术、新知识、新理论、新产品层出不穷,加速了信息资源的新陈代谢。信息资源的半衰期指的是某学科领域目前尚在使用的全部信息资源中较新的信息资源出现的年限。据统计不同学科信息资源的半衰期为:生物医学 3 年、冶金学 3.9 年、物理学 4.6 年、化工 4.8 年、社会科学 5 年、机械工程学 5.2 年、生理学 7.2 年、化学 8.1 年、数学 10.5 年、地质学 11.8 年、地理学 16.1 年。由于各国科技发展水平不同,其相应的信息资源寿命也不尽相同。

(四) 电子信息逐渐成为一种主要的信息资源形式并表现出鲜明的特点

计算机技术和情报资源的结合产生了一种新型的信息资源——电子信息,这种信息资源主要以数字形式存储,用户利用计算机或互联网进行检索和传输,并通过书目、索引、文摘、全文数据库、多媒体信息、电子期刊等形式为用户提供服务,加快了信息资源的流通和共享。另外,网络信息呈全球化分布结构,信息资源分别存储在不同国别、不同区域的服务器上,不同的服务器采用的数据结构和操作系统也不同,从而使网络信息资源的呈现方式从静态的文本格式发展到动态的多模式的方式。它的组织结构完全超越了传统的信息组织方式,具有一定的复杂性。

三、信息资源的类型

信息资源的表现形式多种多样,其类型的划分并没有固定的标准,主要取决于人们分析问题的不同需要,因此可以从不同需要出发按照多个角度对信息资源类型进行划分。

(一) 根据信息资源的记录载体划分

目前信息资源载体的主流形式大体可以划分为两类:印刷型信息资源和电子型信息资源。而电子型信息资源又可以根据符号编码系统和制作生产过程进行进一步划分。印刷型信息资源是信息的传统记录形式,它包括石印、油印、铅印、胶印和光电印刷。电子型信息资源是伴随计算机技术和网络技术发展而产生的,它是以数字信息处理技术为核心的一种信息记录形式。

根据数字信息记录的不同符号编码系统,电子信息资源可以进一步地划分为三种类型:文本类电子信息资源、图形类电子信息资源和多媒体类电子信息资源。其中,文本类电子信息资源是通过语言、文字输入后产生的电子信息资源,它主要借助计算机电子文本处理软件产生的文件保存格式,其主要格式包括 TXT 格式、RTF 格式、HTML 格式、PS 格式、DOC

格式、PDF 格式、PPT 格式、XLS 格式等;图形类电子信息资源的文件保存格式主要有 BMP 格式、JPG 格式、GIF 格式、TIFF 格式、CDR 格式、DXF 格式、DWF 格式、KML 格式、SWF 格式等;而多媒体类电子信息资源的主流格式包括 CD 格式、WAVE 格式、MIDI 格式、WMA 格式、MPEG 格式、Real 格式、WMV 格式、QuickTime 格式、MKV 格式等。

另外,根据数字信息记录的不同制作、生产过程,电子信息资源还可以进一步地划分为原生电子信息资源、次生电子信息资源和数字化电子信息资源三类。目前,在国内外的商业数据库中,大多数是既有原生电子信息资源,也有次生电子信息资源、数字化电子信息资源。由于受到版权许可证的限制,只有数字信息出版商和少部分数字信息服务商能够制作原生电子信息资源,绝大多数数字信息集成商和数字信息服务商只能制作次生电子信息资源。

电子信息资源的出版、制作、服务商基于自身经济利益和著作权保护的目,在信息加工、处理技术上,有的采用国际通用标准的文件存储格式,有的则采用独家标准的文件存储格式。当前,国内电子期刊、数字图书的文件存储格式有 PDF、PDG、CAJ、VIP、CEB 等,其中 PDF 为国际通用格式,PDG、CAJ、VIP、CEB 分别为国内数字集成商超星、清华同方、维普资讯和方正 Apabi 采用的独家标准存储格式。

(二) 根据信息资源的传递加工层次划分

根据信息资源的传递加工层次可以划分为零次信息、一次信息、二次信息和三次信息。其中,零次信息是一切信息产生的源信息,即客观存在于社会生活中,通过人的视觉、听觉等形成的言语、神情、气氛等表象形式,其主体是口头信息及行为表现,与零次信息相对应的是零次文献,是指非正式出版或非正式渠道交流的最原始资料,如手稿、个人通信、原始记录、演讲及实验记录等是最基本的信息源。一次信息也称原始信息,它是作者依据本人的科研和工作成果而形成的信息,是对作者学术思想、研究成果、技术发明、文艺创作等的正式记载,是科研成果的一种主要表述方式。此类信息资源的印刷形式主要包括图书、期刊和报纸、研究报告、学位论文、会议论文、专利说明书、技术标准、产品样本等;电子形式包括事实数据库、电子期刊、电子图书、电子预印本和发布一次信息的正式学术网站等。一次信息具有创造性、原始性和多样性的特点。

二次信息有时又称检索工具,它指的是根据实际需要,按照一定的科学方法,将特定范围内的分散的一次信息进行筛选、加工、整理,使之有序化而形成的信息。由于它能较为全面地、系统地反映某学科、某专业的文献信息线索,因而是检索和评价一次信息的便捷工具。此类信息的印刷形式有期刊、数据、文摘、题录、索引、论文集、图书馆目录等。其中,二次信息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。二次信息主要有集中性、工具性和系统性的特点。

三次信息是指通过二次信息提供的线索,选用一次信息的内容进行分析、综合、研究后而编成的信息文献,它是围绕某一专题,对各次信息所含知识的深度加工、浓缩和重组。一般包括专题评述、专题调研、动态综述、进展报告、学科年度总结等。此类信息资源的印刷形式和电子形式基本重合,都包括综述、述评、辞典、百科全书、年鉴、标准、数据手册等。三次信息的特点主要表现在综合性、针对性和科学性上。

(三) 根据信息交流的形式划分

从人类社会的发展历程来看,信息交流的形式经历了三种不同的阶段和形态:非正式信

息、半非正式信息和正式交流。因此,可以将信息资源划分为非正式信息资源、半非正式信息资源和正式信息资源三种类型。

1. 非正式信息资源

非正式信息资源是人与人之间直接互动交流而产生的信息,又称零次信息或零次文献。非正式信息资源,基本上是通过面对面谈、电话、演讲、邮件或书信往来、研讨会的论文发表等渠道传播。零次信息是一次信息资源形成前对作者学术思想、研究成果、技术发明、文艺创作等的最原始的记录,此类信息的印刷形式为手抄件、打字件、复印件、讲义、描图、黑板板书,电子形式为内部录音、录像、电子邮件、电子文档等。零次信息的特点是信息来源直接、真实,内容新颖。这类信息在较小的范围内交流、使用、参考,其传播渠道少。非常重要的零次信息资源则作为档案或制作数据备份以长期保存。

2. 半非正式信息资源

半非正式信息资源是指受到一定产权保护但又没有纳入正式出版系统而产生的信息,又称“灰色”信息。部分半非正式信息由于保密需要,不能通过常规的手段获取,只在特定的范围进行传播;部分半非正式信息虽然未公开发表,但又允许在一定范围内传播,如非公开的内部刊物、专题报告、市场调查报告、学位论文、学术会议文献等。半非正式信息具有原始性、不稳定性、隐蔽性、离散型、获取困难、形式灵活多样等特点,它能够真实反映一个国家、一个行业、一个机构发展的动态和水平。而且,这些信息由于传播的范围小,被多次利用的可能性也比较小,在实际社会活动中往往会产生较大的收益。

3. 正式信息资源

正式信息资源是指经过出版机构和同行专家以严格的专业标准进行过评审,受到严格知识产权保护的知识性信息,包括图书、期刊、报纸、专著、工具书、标准、技术报告、音像制品等。正式信息资源必须经过出版机构的资格认证、同行评审、质量控制三个环节,才能进入正式的信息交流渠道。更重要的是,由于出版机构的竞争,导致只有最优秀的知识性信息才能在最优秀的出版物发表。正式信息是人类最主要的文明和知识的总和,具有稳定性、科学性、质量可靠、传播范围广和利用率高的优点,能够充分地反映一个国家的文明发展水平。在大数据时代,重视正式信息资源的寻找和利用,尤其是充分重视那些优质正式信息的获取和吸收,可使学习和研究少走弯路。

第二节 主要信息资源

数字信息资源的出版形式是图书馆和信息服务机构存放、管理和提供服务的一个依据,根据文献出版形式的不同,文献型数字信息资源主要有以下十种类型。

一、图书

图书是最早出现的文献类型之一,至今仍占据科学文献的主导地位。联合国教科文组织在1961年规定:5~49页印刷品为小册子,50页以上的正式印刷品为图书。图书主要有以下优缺点:出版量大,稳定性强,质量较高,保存期长,但编辑出版周期较长,传递知识信息

相对较慢。图书主要包括专著、教科书、工具书、图册、读物和文集等。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)。其中出版社名称^①、出版地或出版国^②、出版时间^③、总页数或页码范围^④、国际标准书号(ISBN)^⑤是辨识图书的主要外部特征;图书辨识的直接关键词^①是“出版(社、者)”,英文词是 Press、Publication(Pub.)、Publisher。例如:

(1) 江永红,主编. 统计学. 合肥^②:中国科学技术大学出版社^①,2002^③,239 页^④, ISBN 7-312-01363-5^⑤

(2) C. Granell, J. Poveda, M. Gould. Incremental Composition of Geographic Web Services: An Emergency Management Context.. F. Toppen, P. Pastracos (eds.): Proc. of the 7th Conference on Geographic Information Science (AGILE 2004^③) Heraklion^②, Creta (Grecia)^②. Abril 2004^③, Crete University Press^①, pp. 343-348^④ (ISBN: 960-524-176-5^⑤)

(3) Javier Nogueras-Iso, F. J. Zarazaga-Soria, P. R. Muro-Medrano. Geographic Information Metadata For Spatial Data Infrastructures—Resources, Interoperability, Information Retrieval. Springer Verlag^①, Germany^②, ISBN: 3-540-24464-6^⑤, 264 pages^④, March 2005^③

其中,ISBN 号是国际标准书号(International Standard Book Number)的简称,它是国际标准化组织于 1972 年、2005 年公布的一项国际通用的出版物统一编号方法。2007 年 1 月 1 日以前的 ISBN 号由 10 位数字组成,以后由 13 位组成,共分五段:① 前缀 978 或 979,为 3 位 EAN(欧洲商品编号),代表图书。② 地区号,又叫组号(Group Identifier):代表出版者的国家、地理区域、语种等;我国的组号为“7”。③ 出版者号(Publisher Identifier),代表组内所属的一个具体出版者(出版社、出版公司等)。④ 书序号(Title Identifier):代表某出版社所出版的一种具体出版物的书名。⑤ 校验码(Check Digit):ISBN 号的最后一位数字,用以检查 ISBN 号转录过程中有无差错。例如,《Frontpage 网页设计》一书的 ISBN 号是:978-7-118-01984-4,其中:“978”为欧洲商品编号中图书;“7”代表中国(中文);“118”代表国防工业出版社;“01984”代表书名号;“4”代表检验码。

二、期刊

期刊又称杂志、连续出版物,指有固定名称、统一开本、有编号或年月标志、定期或不定期连续出版、每期内容不重复并由多名责任者撰写不同文章的出版物。与图书相比,期刊的特点是出版周期较短、内容新颖、学术性强、信息量大并能及时反映各学科发展的最新动向和科学研究的最新成果。期刊是随着近代科学的发展而产生的,其发展十分迅速,在科技活动中起着非常重要的作用,是科技人员进行信息交流的正式公开的工具,被称为“整个科学史上最成功的无处不在的科学信息载体”。

期刊的类型有五大类,并且有自己的常用冠名:

(1) 学术性、技术性刊物,是科技期刊的核心部分。刊名一般为:Acta(学报)、Journal(杂志)、Annals(年报)、Chronicle(纪事)、Annual(年刊)、Bulletin(通报)、Transactions(汇刊)、Proceedings(会刊)、Review(评论)、Progress 或 Advances in ... (进展)等。

(2) 快报性刊物。刊名一般为:Communication(通讯)、Letters(快报)、Bulletin 等。

(3) 消息性(Newsy Journals)刊物。多数是新闻,刊名一般为:News(新闻)、News Letters(新闻快报)等。

(4) 资料性(Data Journals)刊物。刊名一般为:Data、Event 等。

(5) 检索性刊物。刊名一般为:Index、Abstract 等。

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.)、年月号、起止页号、国际标准刊号(ISSN)。其中:卷号(Vol.)^①、期号(No.)^②、年月号^③、起止页号^④、国际标准刊号(ISSN)^⑤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^⑥。例如:

(1) 王玲. 基于知识发现的生物信息学. 生物工程进展^⑥, 2000^③, Vol. 20^①, No. 3^②: 27 - 29^④, ISSN 1003 - 3565^⑤

(2) Guterman, L. The Promise and Peril of "Open Access". Chronicle^⑥ of Higher Education, January 30, 2004^③, 50^①(21^②): A10 - 12, A14^④, ISSN: 0009 - 5982^⑤

(3) Poveda, J. and M. Gould. The BQV algorithm for point-in-polygon determination. GESTS International Transactions^⑥ on Computer Science and Engineering, Vol. 6^①, No. 1^②, pages: 207 - 218^④. Jan. 2005^③

其中,国际标准刊号(ISSN)由 8 位数字分两段组成,如 1000~0135,前 7 位是期刊代号,末位是校验号。我国正式出版的期刊都有国内统一刊号(CN),它由地区号、报刊登记号和《中图法》分类号组成,如 CN11 - 2257/G3。地区号依《中华人民共和国行政区划编码表(GB 2260 - 82)》取前两位,如北京为 11、天津为 12、上海为 31、辽宁为 21、吉林为 22、四川为 51 等。图书馆一般把当年的期刊称为“现刊”,当年以前的期刊称为“过刊”,相应地也有“现刊阅览室”和“过刊阅览室”的设置和称呼。

三、会议文献

会议文献是指在学术会议上宣读或书面交流的报告、论文、会议记录、会议纪要等有关资料,是科学交流的一条重要渠道。它往往反映出科学技术的发展趋势,其特点是与最新成果的间隔时间短,但其内容与期刊相比可能不太成熟。会议文献具有水平高、针对性强和发表快的特点,很多论文往往在正式发表前先提交学术会议进行交流。全世界每年召开的学术会议超过上万次,正式发行的各种专业论文集、会议录近 4000 种,约有 50% 的会议只在会前印发论文、论文摘要或论文目录等会前文献,不出版会后论文集、会议录。

会议和会议文献常用的主要名称有 Conference(大会)、Meeting(小型会议)、Symposium(讨论会)、Seminar(研讨会)、Proceeding(会议录)、Paper(单篇论文)、Transaction(汇报)等。英文常用 Transactions 表示会议上发表的汇编论文;Proceedings、Symposium 表示会议的记录和会后整理出版的会议论文录、论文集。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、会议年月日、论文在会议论文集中起止页号、会议论文号。其中,会议名称或会议论文集名称^①、会议地或主办国^②、会议年月日^③、论文在会议论文集中起止页号^④、会议论文编号^⑤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^⑥。例如:

(1) X. Gai, S. Sun, M. F. Wheeler and H. Klie, A time stepping scheme for coupled reservoir flow and geomechanics on nonmatching grids, In: Proceedings^⑥ of the 2005 SPE Annual Technical Conference and Exhibition^①, Dallas, Texas, USA^②, October 9-12, 2005^③. (SPE 97054^⑤)

(2) S. Sun and M. F. Wheeler, Adaptive discontinuous Galerkin methods for coupled diffusion—and advection-dominated transport phenomena, In: Proceedings^⑥ of the 3rd International Conference on Computing^①, Communication and Control Technologies (CCCT'05), Volume I, ISBN: 980-6560-46-9, H.-W. Chu, M. J. Savoie, and B. Sanchez, Eds., pp. 130-135^④, Austin, Texas^②, July 24~27, 2005^③.

(3) R. Hedges, B. Loewe, and C. Morrone. Parallel File System Testing for the Lunatic Fringe: the care and feeding of restless I/O Power Users. Twenty-Second IEEE/Thirteenth NASA Goddard Conference^⑥ on Mass Storage Systems and Technologies (MSST 2005)^①, Monterey, California, USA^②. 2005^③, Pages. 3-17^④

四、学位论文

学位论文指高等学校、科研机构所培养的硕士生、博士生为取得学位资格而呈递的毕业论文,是带有一定独创性的原始文献,具有较高的学术水平和学术价值。世界各国都非常重视学位论文的收藏和流通,并出版目录和索引。

学位论文是非卖品,也不公开发行,通常只在学位授予单位的图书馆和按国家规定接受呈缴本的图书馆保存有副本,故学位论文的收集与利用不如其他类型的文献方便。例如,在中国,国家科技文献中心(NSTL)、中国科技信息研究所、万方数据、CNKI(同方)都集中收藏和报道国内各学位授予单位的博/硕士学位论文。在美国,由国际大学缩微品公司(University Microfilms International, UMI)负责收藏和报道全美的博士、硕士论文,并收集报道其他国家的学位论文。在英国,由英国国家图书馆(不列颠图书馆)负责收藏和报道本国的学位论文。在日本,国立大学的学位论文由日本国立图书馆统一收藏,私立大学的学位论文则收藏在学位授予单位的图书馆中。

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^①是“学位论文”和“学位名称”,英文词是 Doctoral Dissertation 和 MS、M. B. A.、Ph. D.、Eng. D、D. S. 等。例如:

(1) Chang, D. H. (2000). Knowledge, culture, and identity: American influence on the development of Library and Information Science in South Korea. Unpublished doctoral dissertation^①, Univ. of Texas at Austin

(2) 2005, Tsai, Hwai. Pricing Discrete Double-Barrier Options with the Quadrature Method. MBA thesis^①.

(3) Allen, B., Learning Body Shape Models from Real-World Data, Ph. D. Thesis^①, 2005.

五、标准文献

标准文献是经权威主管机构批准,采用文件形式或规定基本单位(物理常数)这两种形

式固定下来、以文件形式出现的法定性文献。标准文献以特定形式对重复性事物和概念做统一规定,是科学、技术和社会实践中非常重要的信息资源。

按标准文献内容划分有基础标准、产品标准、方法标准、安全卫生标准等;按成熟程度划分有法定标准、推荐标准、试行标准;按使用范围划分有国际标准、区域标准、国家标准、行业标准、企业标准等。国际标准化机构中最重要、影响最大的是 1947 年成立的国际标准化组织 (ISO)和 1906 年成立的国际电工委员会 (IEC),它们制定或批准的标准具有广泛的国际影响。

标准文献都有标准号,它通常由国别(组织)代码+顺序号+年代组成,如 ISO 3297-1986。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),如 GB 18187-2000、GB/T 2662-1999;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、SY 表示石油行业标准;企业标准编号:Q/省、市简称+企业名代码+年份。国际常用标准文献的代码,如表 1.1 所示。

表 1.1 国际常用标准文献的代码

ISO	国际标准化组织	DIN	德国国家标准
IEC	国际电工委员会	ГОСТ	俄罗斯国家标准
ANSI	美国国家标准	GB	中国国家标准
BS	英国国家标准	ASME	美国机械工程师协会标准
CEN	欧洲标准化委员会	ASTM	美国材料和实验标准
CENELEC	欧洲电子技术标准委员会	API	美国石油协会标准
JIS	日本工业标准	IEEE	美国电气与电子标准
NF	法国国家标准	ITU	国际电信联盟标准

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^①是“标准”(Standard)与“标准号”。例如:

- (1) 评定水在烃和脂肪酯润滑剂中溶解度的标准^①试验方法,ASTM D 4056-2001^①
- (2) 信息技术自动识别与数据采集技术条码码制规范^① 交插二五条码,GB/T 16829-2003^①
- (3) Recommended Practice Standard^① Procedure for Field Testing Water-Based Drilling Fluids. Second Edition, Addendum, 1 May 2000. API RP 13B-1^①
- (4) NACE Standard^① RP0775 - 2005^①. Preparation, Installation, Analysis, and Interpretation of Corrosion Coupons in Oilfield Operations. 2005-04-07.

六、专利文献

专利是受专利法保护的技术发明,并独占其制造、使用和销售的权利。专利文献是实行专利的国家、地区及国际专利组织在审批专利过程中产生的官方文件及出版物,包括专利说明书、专利权利要求书、专利公报、专利分类表、专利检索工具等。专利是与产业活动密切相关的,因而其实用性非常强。

根据专利权种类及其应用情况,其类型有发明专利、实用新型专利和外观设计专利等。

发明专利是指对产品、方法提出的新方案或对原有产品、方法提出的改进。实用新型专利是指对产品的形状、构造或组合提出的新的实用技术方案。外观设计专利是指对产品的形状、图案、配色或其结合做出的新颖设计,一般不出版说明书。专利文献辨识的直接关键词^①是“专利号”(包括申请号、公开号)。例如:

(1) 一种消除注气驱油过程中固溶物沉淀伤害的方法。CN200310103591.7^①(申请号),CN1616794^①(公开号)。

(2) Shmakova-Lindeman, Olga E. Paraffin inhibitors. US 20050215437^①

(3) Miklos, H. Method for the treatment and prevention of asphaltene-paraffin-wax precipitates in oil-wells, wellheads and pipelines by the use of biocolloid suspensions. EP1552106^①; WO2004031536(A1)^①; CA2501359(A1)^①; AU2003274389^①

七、报纸

报纸在国外被称为新闻报纸,是专门刊载新闻报道和时事评论为主的定期出版物,但有些报纸也刊载学术论文和其他信息。报纸具有时事性、时效性、普及性、大众性的特点,而且出版周期短,传递信息迅速,传播范围广,也是不容忽视的文献信息资源。

八、政府出版物

政府出版物是指国家政府机关颁布的各种文件,如政府公报、政府会议文件、法律汇编、条约集、公告等,国家党政机关为开展日常工作而印刷的文件、参考资料也属于此类。政府出版物具有正式性、权威性的特点,它与国际国内政治经济形势密切相关,是科技工作者十分重视的文献信息资源。

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,其中美国政府出版局是世界上最大的政府出版机构。美国政府出版物可通过美国政府出版物目录月报(<http://www.access.gpo.gov>)、美国政府报告通报与索引(Government Reports Announcement & Index, GRAI)和国家科技文献中心利用。欧洲国家政府出版物可通过欧共体信息网政府联机信息服务系统(<http://europa.eu.int/index.htm>)和国家科技文献中心利用。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,部分政府出版物也发布在人民网、新华网上。

九、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。科技报告一般都编有号码,供识别报告本身及其发行机构,其特点是反映新技术、新学科较快,内容专深,数据可靠,保密性强,具有独创性。

科技报告一般单独成册,有专门的编号。较著名的有美国国家技术信息服务处(NTIS)出版的四大报告:PB报告、美国国防技术信息中心(DTIC)出版的AD报告、美国国家航空航天局出版的NASA报告、美国能源部出版的DOE报告。

科技报告的种类很多,按照报告形式可分为报告书、技术札记、备忘录、论文、通报、技术译丛、特种出版物;按照报告时间可分为初期报告、进展报告、中间报告、终结报告;按照报告