

信息资源管理系列丛书

信息检索与利用

魏联华 孙艳美  主 编
陈兆生 李雁行  副主编

XINXI JIANSUO YU LIYONG



知识产权出版社

INTELLECTUAL PROPERTY PUBLISHING HOUSE

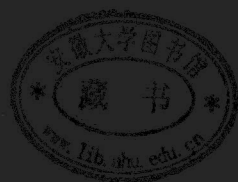
全国百佳图书出版单位

信息资源管理系列丛书

信息检索与利用

魏联华 孙艳美 □ 主 编
陈兆生 李雁行 □ 副主编

XINXI JIANSUO YU LIYONG



内容提要

本教材针对高职高专学生特点, 注意培养学生信息素质, 即信息获取能力, 其主要指导思想是淡化理论注重实践能力培养, 但理论又贯穿始终, 同时注意结合实例, 使文献检索与信息素质教育有机地结合起来, 以期学生的能力和素质得到更全面的发展。本书共分七章主要包括检索基础知识、馆藏书目和电子图书检索、网络信息资源检索、中外文数据库检索、特种文献检索和论文写作等内容。

责任编辑: 许 波

图书在版编目(CIP)数据

信息检索与利用/魏联华主编; 孙艳美, 陈兆生, 李雁行编. —北京: 知识产权出版社, 2011. 4

(信息资源管理系列丛书)

ISBN 978-7-5130-0392-6

I. ①信… II. ①魏…②孙…③陈…④李… III. ①情报检索—高等学校: 技术学校—教材
IV. ①G252. 7

中国版本图书馆 CIP 数据核字 (2011) 第 021575 号

信息检索与利用

魏联华 孙艳美 主编

出版发行: 知识产权出版社

社 址: 北京市海淀区马甸南村 1 号

网 址: <http://www.ipph.cn>

发行电话: 010-82000860 转 8101/8102

责编电话: 010-82000860 转 8127

印 刷: 保定市中国画美凯印刷有限公司

开 本: 787mm×960mm 1/16

版 次: 2011 年 4 月第 1 版

字 数: 254 千字

ISBN 978-7-5130-0392-6/G·379(3306)

邮 编: 100088

邮 箱: bjb@cnipr.com

传 真: 010-82005070/82000893

责编邮箱: xubo@cnipr.com

经 销: 新华书店及相关销售网点

印 张: 13. 25

印 次: 2011 年 4 月第 1 次印刷

定 价: 26. 00 元

出版权专有 侵权必究

如有印装质量问题, 本社负责调换。

编 委 会

主 编 魏联华 孙艳美

副主编 陈兆生 李雁行

参 编 唐 静 李璐璐 刘桂丽
谢 轩 宗永升

前 言

1984年教育部继《印发〈关于在高等学校开设《文献检索与利用》课的意见〉的通知》，1992年5月原国家教委又颁布了《文献检索课教学基本要求》，对文献检索课的课程性质、教学目的等作了细致而全面的规定，从而使文献检索课从形式到内容更加规范化、系统化。文献检索课是普通高等学校大学生必修课之一，高职院校培养的是社会急需的应用型、技能型人才，在现代科技进步和社会经济发展越来越依赖于信息技术及信息资源的开发与利用的今天，高职院校的学生更应具备从文献和数据库中获取知识信息的能力，不断提高自学能力和科技能力。

《信息检索与利用》课就其实质来说，是一门方法课，即讲授检索和利用各类信息资源的方法，主要是培养学生的信息素养和检索技能。本教材针对高职高专学生特点，注意培养学生信息素质即信息获取能力，因此，教材编写者紧紧围绕信息获取能力这一核心组织教材内容，并注意结合实例，使文献检索与信息素质教育有机地结合起来，以期学生的能力和素质得到更全面的发展。

本教材共分7章，主要指导思想是淡化理论注重实践能力培养，但理论又贯穿始终。每一章的开篇有学习目标，最后附有习题。本书主要包括信息检索基础知识、馆藏书目和电子图书数据库检索、网络信息资源检索、中外文数据库检索、特种文献信息资源检索和毕业论文写作等内容。具体编写分工如下：第1章由孙艳美、魏联华编写，第2章由孙艳美、陈兆生编写，第3章由李雁行、魏联华编写，第4章由李璐璐编写，第5章由李璐璐、陈兆生编写，第6章由唐静编写，第7章由刘桂丽编写，谢轩、宗永升对部分章节进行了补充和审校工作。

在本书的编写过程中也得到了秦皇岛职业技术学院图书馆和相关部门同仁的大力支持，在此表示衷心的感谢，同时也借鉴了大量国内专家和学者的著作，在此一并表示真诚的谢意。囿于编者的学识和水平，书中不免有疏漏和不足之处，恳请读者批评指正。

编 者

二〇一一年元月

目 录

第1章 信息检索基础知识	(1)
1.1 信息概述	(1)
1.1.1 信息及其相关概念	(1)
1.1.2 文献的级次及其主要类型	(3)
1.2 信息检索的概念和类型	(6)
1.2.1 信息检索的含义	(6)
1.2.2 信息检索的类型	(7)
1.3 信息检索语言	(8)
1.3.1 分类语言和分类法	(9)
1.3.2 主题语言与主题法	(12)
1.4 信息检索途径与方法	(13)
1.4.1 信息检索途径	(13)
1.4.2 信息检索方法	(14)
1.4.3 计算机检索技术	(15)
1.4.4 检索步骤	(17)
1.4.5 检索效果的评价	(19)
习题	(20)
第2章 馆藏书目及电子图书数据库检索	(21)
2.1 馆藏书目查询	(21)
2.1.1 汇文书目检索	(22)
2.1.2 我的图书馆	(23)
2.2 超星数字图书馆	(24)
2.2.1 数据库简介	(24)
2.2.2 检索方法	(25)
2.2.3 超星阅读器	(28)
2.3 方正 (APABI) 数字图书馆	(30)
2.3.1 数据库简介	(30)
2.3.2 借阅流程	(30)



2.3.3 Apabi Reader 阅读器	(33)
2.4 九羽数字图书馆	(36)
2.5 五车数字图书馆	(36)
2.6 免费中文电子图书	(37)
习题	(37)
第3章 网络信息资源检索	(39)
3.1 搜索引擎	(39)
3.1.1 搜索引擎及其分类	(39)
3.1.2 常用全文搜索引擎	(40)
3.1.3 常用目录型搜索引擎	(49)
3.1.4 常用元搜索引擎	(52)
3.1.5 专题型搜索引擎	(54)
3.2 事实和数据网络检索	(56)
3.2.1 字典、词典类工具书网站	(57)
3.2.2 百科全书类工具书网站	(59)
3.2.3 年鉴类工具书网站	(60)
3.2.4 名录类工具书网站	(61)
3.3 学科专业资源特色检索	(62)
3.3.1 经济管理类信息检索	(62)
3.3.2 旅游类信息检索	(64)
3.3.3 艺术类信息检索	(66)
3.3.4 计算机、机电类信息检索	(67)
3.4 开放存取资源检索	(68)
3.4.1 开放存取概述	(68)
3.4.2 开放存取资源	(69)
习题	(73)
第4章 通用外文数据库	(75)
4.1 三大检索系统简介——SCI、EI、ISTP	(75)
4.1.1 《科学引文索引》(SCI)	(75)
4.1.2 《工程索引》(EI)	(87)
4.1.3 《科学技术会议录索引》(ISTP)	(95)
4.2 Elsevier SDOS 数据库	(99)
4.2.1 Elsevier 和 Science Direct 简介	(99)
4.2.2 SDOS 数据库的使用	(100)
习题	(105)

第5章 常用中文数据库检索与利用	(107)
5.1 中国知网 (CNKI) 数据库	(107)
5.1.1 CNKI 简介	(107)
5.1.2 CNKI 平台的使用	(108)
5.1.3 知网节的使用	(117)
5.1.4 CAJ 浏览器的使用技巧	(118)
5.1.5 工具书及知识元检索	(120)
5.1.6 检索实例	(121)
5.2 维普数据库	(123)
5.2.1 维普信息资源系统概述	(123)
5.2.2 检索方法	(123)
5.3 万方数据库	(127)
5.3.1 万方数字资源简介	(127)
5.3.2 检索方法	(128)
习题	(131)
第6章 特种文献信息检索	(133)
6.1 标准与标准文献	(133)
6.1.1 标准的基础知识	(133)
6.1.2 标准文献	(136)
6.1.3 标准文献检索	(137)
6.2 专利与专利文献	(152)
6.2.1 专利	(152)
6.2.2 专利文献	(155)
6.2.3 专利文献信息检索	(161)
6.3 会议文献的检索	(174)
6.3.1 会议文献概述	(174)
6.3.2 会议文献的检索工具	(175)
6.3.3 网上会议文献资源	(176)
习题	(178)
第7章 毕业论文写作	(180)
7.1 毕业论文写作概述	(180)
7.1.1 毕业论文写作的意义	(180)
7.1.2 学术论文的分类	(180)
7.1.3 毕业论文的基本要求	(182)
7.1.4 毕业论文的构成项目与写作规范	(183)



7.2 毕业论文写作步骤	(189)
7.2.1 课题资料的查询	(189)
7.2.2 文献综述	(190)
7.2.3 毕业论文的选题	(190)
7.2.4 毕业论文的开题报告	(191)
7.2.5 论文资料的搜集与整理	(191)
7.2.6 立意谋篇	(193)
7.2.7 论文初稿撰写与修改	(195)
习题.....	(198)
参考文献.....	(199)

第 1 章 信息检索基础知识

■ 学习目标

通过本章的学习，能够了解信息检索的基础知识，包括信息与信息检索的概念、检索语言等，能够有效地利用信息检索的基本方法进行信息检索，理解有关文献、信息检索、检索语言具体的种类及特点，重点掌握信息检索的基本方法、途径和步骤，为学习从海量的网络信息中检索有用知识和从事研究打下坚实的基础。

1.1 信息概述

1.1.1 信息及其相关概念

信息，英语“information”，来源于拉丁语，我国大陆也译为情报或信息，在中国香港、台湾地区被译为“资讯”。在我国，信息一词古已有之，据考证，最早使用“信息”一词的人为唐朝崔备（唐建中二年公元 781 年进士及第），他在《清溪路中寄诸公》中有“别来无信息，可谓井瓶深”。其后杜牧诗《寄远》中有云：“塞外音书无信息，道旁车马走尘埃”。唐诗中的信息大多指音讯、消息。信息作为一个科学概念，直到 20 世纪中叶才被提出，是指通信系统传输和处理的对象。随着社会信息量的剧增，信息概念逐步运用到各种领域。

1. 信息（Information）

什么是信息？它有着多种多样的含义，一直以来，在不同领域人们赋予它不同的定义，具有代表性的主要有以下几种观点：

美国信息论创始人申农（C. E. Shannon），在 1948 年写的《通信的数学理论》（《*Mathematical Theory of Communication*》）中指出一个系统所接收的信息是“能够用来消除不确定性的东西”。

美国控制论奠基人诺伯特·维纳（N. Wiener），在他的《控制论或关于在动物和机器中控制和通讯的科学》（《*Cybernetics or Control and Communication in the Animal and the Machine*》）中认为，信息是“我们在适应外部世界和控制外部世



界的过程中同外部世界交换的内容的名称”。

《辞海》(1999 年版), 将信息定义为音讯、消息; 通讯系统传输的对象。第一个释义是传统的信息含义, 第二个则带有明显的现代技术特征。

总之, 信息无时无刻不存在于我们的生活和学习中, 信息的爆炸式增长速度和利用已成为科技进步的重要标志之一。综合上述观点, 我们采用我国信息学家钟义信在《中国大百科全书》图书馆学情报学档案学卷中的说法, 将信息定义为: 信息是与物质、能量并立的现代社会三大支柱之一。按照狭义的理解, 信息是用来消除不确定性的东西。按照广义的理解, 又有两种认识: 从本体论意义上说, 信息泛指一切事物(物质的、精神的)运动的状态和运动的方式, 包括事物内部结构的状态和方式以及外部联系的状态和方式; 从认识论意义上说, 信息是关于事物运动状态和运动方式的反映。

2. 知识 (Knowledge)

《辞海》(1999 年版) 对知识的解释为: 知识是人类认识的成果和结晶。知识在牛津英语字典里解释为“通过经验获得的认识, 是人的信息范畴, 是理论或实践上的理解, 是对已知事实的总结”。

经济合作与发展组织 (OECD) 的报告《以知识为基础的经济》(1996) 给出了知识的 4W 概念, 即 (know - what) 知道是什么, (know - why) 知道为什么, (know - how) 知道怎么做, (know - who) 知道是谁。我国的知识经济专家袁正光教授对此是这样解释的: “知道是什么的知识, 指关于事实方面的知识, 如纽约有多少人口, 中国有多大面积; 知道为什么的知识, 指原理和规律方面的知识, 如牛顿三大定律、市场机制、供求规律等; 知道怎么做的知识, 指操作的能力, 包括技术、技能、技巧和诀窍等; 知道是谁的知识, 包括了特定关系的形成, 以便可能接触有关专家并有效地利用他们的知识, 也就是关于管理的知识和能力。”

知识是人们在改造世界的实践中不断接受客观事物发出的信息所获得的认识, 是能够指导思考、行为和交流的正确和真实的观察、经验和过程的总结。包含真理和信念, 观点和概念, 判断和预期, 方法和诀窍等内容, 也可以理解为用于解决问题的结构化信息。

可见知识这一概念与信息的关系最密切。知识是信息的一部分, 是一种特定的信息, “信息”的范畴要远远大于“知识”。知识是人类在改造客观世界的实践中所获得的认识和经验的总和。可以说知识的产生来源于信息, 人们不仅能通过信息感知世界、认识世界和改造世界, 而且能将获得的信息转变成知识作为认识和改造世界的武器。人们把信息转化为知识, 再把知识转化为智慧, 在不断的历史进程和科学进步中, 这是一个动态过程, 是一种开拓过程。反过来, 智慧又转化为新知识, 新知识又转化为新信息, 如此不断循环往复, 又不断创新, 推动

着历史前进。可以说一部人类文明发展史，就是在各种活动中知识的创造——传递——再创造并不断积累和提升的螺旋上升的历史。

3. 文献 (Document)

文献一词，在我国最早出现于《论语·八佾》：“子曰：夏礼吾能言之，杞不足征也；殷礼吾能言之，宋不足征也。文献不足故也。足，则吾能征之矣”。南宋朱熹在《四书章句集注》中对文献的解释为“文，典籍也，献，贤也。”发展到现在，“贤”的意义逐渐取消了。

对文献的定义，按照中华人民共和国国家标准《文献著录总则》(GB/T 3792.1—2009)规定：“文献是记录有知识的一切载体”。具体说是用文字、图形、符号、声频、视频等技术手段记录人类知识的一切载体。从这里可以看出，文献具有三要素：一是有知识信息内容；二是有负载知识的载体材料；三是有记录知识的符号和技术。一般而言，只要记载有知识的信息，不管是纸张或其他实物，无论是否被整理或发表，我们都称之为“文献”。

知识来源于信息，是理性化、优化和系统化了的信息；人们将获得的信息经过大脑加工形成知识，知识被记录在载体上形成文献。文献与知识既是不同的两个概念，二者又有密切的联系。文献包含知识的内容，而知识只能依赖物质载体才能构成文献。文献经过传递、交流、应用又产生信息，如此往复。信息、知识、文献的关系如图1-1所示。

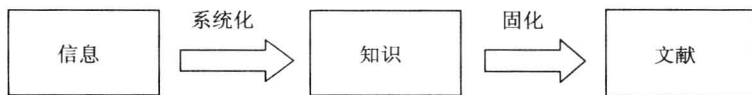


图 1-1 信息、知识、文献之间的关系

1.1.2 文献的级次及其主要类型

1. 文献的级次

文献的级次是根据对文献信息的加工深度来划分的，一般分为4个级次：

(1) 零次文献。是指未经正式出版发行或未进入社会交流的最原始的文献，如私人笔记、底稿、手稿、个人信件、考察记录、实验稿、原始统计数字、技术档案、电子论坛等。这类文献除作者和特定人员外，其他社会成员很难获得和利用。其主要特点是客观性、零散性、内容新颖但不成熟、不公开交流、难以获得。

(2) 一次文献。也称原始文献，是作者本人直接从生产科研和社会活动等实践中产生的研究成果，并在社会上公开交流和发表的文献，如专著、期刊论文、学位论文、会议论文、科技报告、专利说明、技术标准等。这类文献是人们学习参考的最基本的文献类型，是大学生进行信息检索和利用的主要对象，一般会议论文内容更新一些。这类文献的主要特点是内容新颖丰富、真实、具体、富



有创造性,参考价值较大,但也有数量庞大且较分散的缺点。

(3) 二次文献。也称检索性文献,是指对一次文献进行收集、提炼和加工,使之成为系统、有序的文献。这种文献大多被编制成多种检索途径的检索工具,如文摘、索引、题录、书目。在图书馆检索图书时我们经常使用的馆藏书目就是典型的二次文献。这类文献是查找一次文献的工具或手段,可以帮人们在短时间内找到针对性较强的且数量很多的文献。这类文献的主要特点是浓缩性、汇集性、有序性、工具性。

(4) 三次文献。也称参考性文献,是指利用二次文献对检索出的一次文献进行综合分析、研究和评述而形成的文献。三次文献可再分为文献型和数据型。文献型如专题述评、动态综述、进展报告之类;数据型如字典、词典、手册、百科全书、书目之书目等。其中的百科全书是典型的三次文献。这类文献直接来源于一次文献和二次文献但又高于它们,是人们掌握信息源的主要资料。在大学生的平时学习和研究过程中,三次文献是信息检索和利用的对象,也可直接把其作为检索文献,它主要提供各类事实和数据。其主要特点是综合性、参考性。

从零次文献到三次文献,是一个由分散到集中,由无序到有序,由广博到精深的对知识信息进行不同程度加工的过程。一般而言,一次文献是基础,是检索利用的主要对象,二次文献是检索一次文献的检索工具,三次文献则是专门化的高度浓缩的文献信息。不同文献级次如图 1-2 所示。

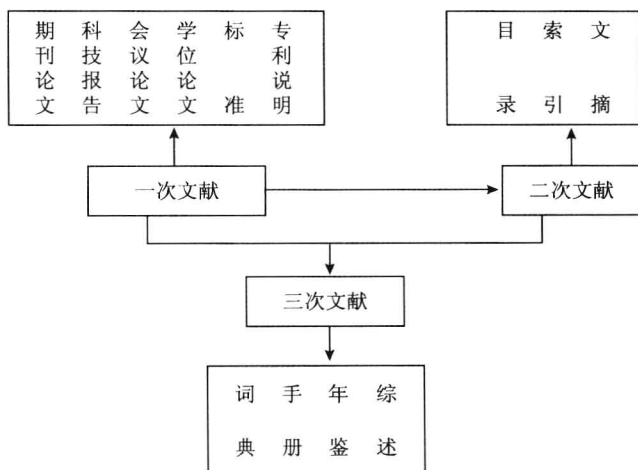


图 1-2 文献级次间关系示意图

2. 文献类型 (按文献的载体形式划分)

按文献的载体形式划分的文献类型包括:刻写型、印刷型、缩微型、声像型、电子型。电子图书、电子期刊、光盘数据库、电视传讯、电子邮件等电子出版物的产生,被认为是人类在知识生产和交流方面的第四次革命。近年来出现的

多媒体或超媒体是一种崭新的文献载体,它采用超文本或超媒体方式,把文字、图片、动画、音乐、语言等多个媒体信息综合起来,在内容表达上具有多样性和直观性,并且有人机交互的友好界面。因此多媒体既属于电子型文献,也属于声像型文献。

3. 文献类型 (按文献的出版形式划分)

国家标准《文献类型与文献载体代码》(GB/T 3469—1983)将文献分成26个类型,即专著、报纸、期刊、会议录、学位论文、科技报告、技术标准、专利文献、产品样本、中译本、手稿、参考工具、检索工具、档案、录音带等。这里将重点介绍几种文献类型。图书、期刊和报纸属于普通文献或传统文献,而其他如专利文献等均属特种文献。

(1) 图书 (Book), 国家标准《情报与文献工作词汇·传统文献》(GB 13143—1991)对图书的解释是:一般不少于49页并构成一个书目单元的文献。图书是文献中最古老、最重要的类型。按写作方式可分为专著、编著、翻译、编译等;按出版卷可分为单卷本、多卷本等;按刊行情况可分为单行本、丛书等;按版次情况可分为初版、重版、修订本等。图书的特点是内容成熟、全面、系统、出版形式较固定,但出版周期长,传播速度较慢。图书主要包括:阅读类图书,如专著、论文集、科普读物等;参考类图书,包括字词典、手册、年鉴、地图、百科全书、名录、指南等。每种图书都有唯一的标识,即国际标准书号——ISBN (International Standard Book Number)。



(2) 连续出版物 (Serials), 包括期刊和报纸。按照国家标准《文后参考文献著录规则》(GB/T 7714—2005)规定,连续出版物是一种载有卷期号或年月顺序号、计划无限期地连续出版发行的出版物。期刊按内容性质分为学术性期刊、通报性期刊、普及性期刊、动态性期刊、检索性期刊;按出版周期分为年刊、月刊、周刊、双月刊、季刊;其特点是连续性、出版周期短,内容新颖、丰富。每种期刊也有国际标准连续出版物号——ISSN (International Standard Serial Number)。

(3) 专利文献 (Patent Document), 广义上的专利文献包括专利说明书、专利公报、专利检索工具以及与专利有关的法律文件等;狭义的专利文献就是专利说明书。专利说明书的内容比较具体,并有附图,通过它可以了解专利的技术内容,因此它是技术人员的重要参考资料。

(4) 技术标准 (Standard), 是标准化组织或有关机构对产品质量、规格、



生产过程以及检验方法等所做的技术规定，是从事制造和生产的共同依据。一个国家的标准资料反映着该国的经济政策、技术政策、生产水平、加工工艺水平、标准化水平以及资源情况等内容。按照审批机构和应用范围划分，技术标准有国际标准、区域性标准、国家标准、部颁标准、企业自订标准等。

(5) 科技报告 (Scientific and Technical Report)，又称研究报告或技术报告，是各国政府系统或科研及生产单位关于某项研究成果的总结报告或研究过程中的阶段性进展报告、成果报告和总结报告。其特点是每份报告为一个专题，单独成册，有机构名称，有统一编号，内容专深而具体，技术含量大，是科技人员的重要参考资料之一。科技报告通常以正式报告、进展报告、技术札记、备忘录等形式出版发行。

(6) 会议文献 (Conference Paper)，又称会议录，是指在国际、国内各种会议上，科技人员宣读、发表或提交的论文、报告等的资料汇编。内容是讨论当前的学术性问题，交流、传递某一专业或领域最新进展的信息资料，有些论文代表着某一学科或专业的国际或国内的研究前沿或研究成果。会议文献能够反映该学科或专业领域的最新水平和发展趋势。

(7) 学位论文 (Dissertation)，是指高等院校或科研院所的学生在取得学士、硕士或博士学位时提交的毕业论文。学位论文是非卖品，一般不公开出版发行，不便搜集，尤其不易全部收藏。

(8) 政府出版物 (Government Publication)，是各国政府及其所属的各个部门发表、出版的文献总称，内容极其广泛，涉及基础科学、应用技术、政治经济、文化体育等。按内容性质可分为两类，一类是行政性文件，如国会记录、政府法令、决议、统计资料等；另一类是科技性文件，包括政府各部门的方针政策、规章制度、研究报告、调查（考察）报告、科普资料以及技术政策等。

(9) 产品样本 (Trade Literature)，又称产品说明书。主要是对定型产品的性能、构造原理、规格、尺寸、用途、使用方法和操作规程、维修方法等所做的具体说明。其特点是形象直观、技术成熟、资料具体、数据可靠。

1.2 信息检索的概念和类型

1.2.1 信息检索的含义 (Information Retrieval)

检索之检为翻阅、考查；检索之索为求取，检索即是通过考查求取所需。“信息检索”一词萌芽于图书馆的参考咨询工作，最早是由美国学者穆尔斯 (Calvin N. Mooers, 1919—1994) 在 1948 年提出，并把它定义为“一种时间性的通讯形式”。这一看法揭示了信息存储与获取两个环节是一种延时性的通讯形式。

信息检索,全名为信息存储与检索,即广义的信息检索,是指将信息按一定的方式组织和储存起来,并能根据用户的需要取出所需特定信息的整个过程。从本质上讲,信息检索包括“存”和“取”两个基本环节。“存”指的是大量信息源的组织化存储;“取”指的是针对各种用户信息需求所进行的分析和检索过程。我们通常所说的信息检索仅仅指信息查找的过程,或是“狭义”的信息检索。

我们说信息检索就是信息用户为处理解决各种问题而查找、识别、获取相关的事实、数据、知识的活动及过程,本质就是信息用户的需求和信息集合的比较与选择,即匹配(match)的过程。从用户需求出发,对一定的信息集合采用一定的技术手段,根据一定的线索与准则找出最相关的信息的过程。

本书中采用叶继元教授对信息检索的定义:信息检索是从大量相关信息中利用人一机系统等各种方法加以有序识别与组织,以便及时找出用户所需部分信息的过程。

1.2.2 信息检索的类型

1. 按检索的手段划分

(1) 手工检索(Manual retrieval)。这是一种相对传统的检索方法,简称“手检”,它是用人工方式查找所需信息的检索方式。手工检索的对象是工具书等书本型的检索工具(包括书本式目录、文摘、索引等),适用于查找早期的纸印刷的书刊文献以及一些不太复杂的检索课题,可以作为机检的补充。它的特点是耗时费力,效率低,查全率较低,很容易造成漏检,但查准率较高,所以现在的大学生和科研人员不到万不得已,很少使用此种检索手段了,最主要的考虑可能是时间成本和地点和条件的限制。

(2) 计算机检索(Computer-based retrieval)。这是现在普遍应用的一种检索方法,简称“机检”,是指把信息及其检索标识转换成电子计算机可以阅读的二进制编码,存储在磁性载体上,由计算机根据程序进行查找和输出。机检的对象是计算机检索系统,针对数据库进行,与手检相比,机检的本质没有改变,变化了的只是信息的媒体形式、存储方式和匹配方法。计算机信息检索的发展与计算机技术、数字化技术、存储技术、网络通信技术等的发展密切相关,从20世纪50年代计算机产生开始至今,应用计算机的信息检索经历了脱机检索、联机检索、光盘数据库检索和网络检索四个阶段。现在我们主要应用的网络检索是利用E-mail、FTP、Telnet、Archie、WAIS、Gopher、Veronica、万维网等检索工具,在Internet等网络上进行信息存取的行为。通过网络接口软件,用户可以在任何一个终端查询各个地区的网上信息资源。

机检与手检的主要区别是机检更要注重检索策略是否得当,概念的选择、数据库的选择、检索策略的制定对于检索结果的获得是至关重要的。计算机检索的



特点是查全率较高,查准率较低,但它还有高效、检索途径多、新颖性、广泛性和不受时空限制等特点,正是这些优势,在校的学生和研究人员应用这种检索手段更为普遍。

2. 按照检索的对象(或检索结果)划分

(1) 文献检索(Document retrieval):这是最基本的一种检索方式,以文献为检索对象的检索,它是指将文献按一定的方式存储起来,然后根据需要从中查出有关课题或主题文献的过程。文献检索可分为书目检索和全文检索。书目检索是以文献检索工具为检索对象,如书目、专题书目、索引和文摘等二次文献。此类检索工具如EI、SCI、《中文期刊数据库》(文摘版)、《全国报刊索引》等。全文检索是以文献所含的全部信息作为检索内容,即检索系统存储的是整篇文章或整部图书。

(2) 事实和数据检索(Fact and Data retrieval)。它是以特定事实和数据(包括图表、公式)等信息作为检索对象,检索结果可以直接被利用。这类数据包括物理性能常数、统计数据、国民生产总值、外汇收支等。此类数据库(检索工具)很多,如《中国企业、公司及产品数据库》、《中国科技名人数据库》、《常用材料性能数据库》、《中国拟建和在建项目数据库》、《中国宏观经济统计分析数据库》、《IMI消费品市场统计分析数据库》等。如果将二者区分开来讲,数据检索是以具有数量性质并以数值形式表示的数据为检索内容的信息检索,或称数值检索;事实检索是以文献中抽取的事项为检索内容的信息检索,或称事项检索。

事实和数据检索与文献检索的主要区别是事实和数据检索是以从文献中提取出来的各种事实、数据为检索对象的一种确定性检索,是信息检索的派生检索,但又是大量的、日常的、有具体对象的检索。如查字或词,查事实概念,查人物,查机构名称和缩写,查地名,查年代日期,查公式、常数、数据、规格等。这些事实和数据是一种确定性的信息检索,检索结果非常明确。而文献检索的检索结果是关于某些问题有哪些相关文献,而不会像事实和数据检索那样得到某些问题的具体检索结果,诸如是哪一个,具体有多少量等。例如文献检索所回答的是诸如“职业院校大学生信息素质教育有哪些文献”这类问题,而事实和数据检索所回答的是诸如“2010年全国高等职业院校在校学生数量是多少”之类的问题。

(3) 声频与视频检索(Audio-visual retrieval)。它是以声音、图像、录像等视听资料为检索对象,其检索结果为图形、图像、视频、音频等的一种检索方式,实现了用户对一些形象化信息的查询与定位。例如,查找老北京、老上海、老天津的照片,查看某地的老建筑图片,查找历届美国总统的就职演讲音频、视频等。

1.3 信息检索语言

信息检索语言是根据信息检索的需要创造出来的一种人工语言,用于描述信