



“十二五”普通高等教育本科国家级规划教材

高等学校图书馆学专业系列教材

第三版

信息检索

主 编 黄如花
副主编 胡永生 程银桂

非外借



WUHAN UNIVERSITY PRESS

武汉大学出版社



“十二五”普通高等教育本科国家级规划教材

高等学校图书馆学专业系列教材

信息检索

第三版

- ▶ 主 编 黄如花
- ▶ 副主编 胡永生 程银桂
- ▶ 编著者 黄如花 胡永生 程银桂 韩 丽 徐 轩
吴 丹 吴文辉 陈振英 杨 眉 耿 哲



WUHAN UNIVERSITY PRESS

武汉大学出版社

图书在版编目(CIP)数据

信息检索/黄如花主编.—3版.—武汉:武汉大学出版社,2019.9
“十二五”普通高等教育本科国家级规划教材
高等学校图书馆学专业系列教材
ISBN 978-7-307-20589-5

I.信… II.黄… III.信息检索—高等学校—教材 IV.G254.9

中国版本图书馆CIP数据核字(2018)第226127号

责任编辑:詹蜜

责任校对:汪欣怡

版式设计:马佳

出版发行:武汉大学出版社 (430072 武昌 珞珈山)

(电子邮箱:cbs22@whu.edu.cn 网址:www.wdp.com.cn)

印刷:湖北金海印务有限公司

开本:720×1000 1/16 印张:20.25 字数:361千字 插页:2

版次:2002年4月第1版 2010年5月第2版

2019年9月第3版 2019年9月第3版第1次印刷

ISBN 978-7-307-20589-5 定价:48.00元

版权所有,不得翻印;凡购买我社的图书,如有质量问题,请与当地图书销售部门联系调换。



主编简介

黄如花，武汉大学二级教授、信息管理学院副院长、博士生导师，教育部新世纪优秀人才。兼任中国图书馆学会学术委员会副主任委员、教育部中国高校人文社会科学文献中心专家咨询组专家、文化部第六次全国公共图书馆评估专家、《中国大百科全书》第三版图书馆学科“信息检索”分支主编。主讲的“信息检索”课程被评为国家精品课程、国家精品资源共享课程、全国大学生优秀素质教育课程和国家精品在线开放课程；主编的《信息检索》教材入选“十一五”“十二五”国家级规划教材。在国内外重要期刊发表信息素养教育相关的论文10余篇，在教育部主管的爱课程网全程独立讲授的“信息检索”MOOC已吸引国内外逾20万学习者，建议并执笔起草的《中国公民信息素养教育提升行动倡议》在2019中国图书馆年会发布。系Government information Quarterly等多个SSCI和iConference等国际会议的同行评审专家。



作者简介

胡永生，研究馆员，武汉大学图书馆咨询部副主任，从事信息素养教育工作20余年，主编或参编多部信息素养类教材，主持并完成了教育部人文社科规划基金等项目。

程银桂，武汉理工大学管理学院副教授，从事管理信息教学与研究20余年，主编多部著作，获省级研究成果奖，成果被ISTP等检索。

韩丽，副研究馆员，北京师范大学图书馆信息服务部主任，长期从事信息素养教育、科技查新、文献计量分析评价等工作。

徐轩，副研究馆员，复旦大学图书馆学科馆员，主要研究方向为信息政策、学科服务、信息素养教育等。

吴丹，武汉大学信息管理学院图书馆学系主任、教授。中组部“万人计划”青年拔尖人才、教育部青年长江学者、中美富布赖特高级研究学者。

吴文辉，法学硕士，中南财经政法大学副研究馆员，从事法学文献检索教学与研究近20年，主要研究方向为法学文献学、法学文献检索。

陈振英，副研究馆员，浙江大学图书馆情报分析中心主任，主要研究方向为科技评估、文献计量与评价。

杨眉，理学博士，上海交通大学图书馆副研究馆员，长期从事科技查新、科学计量学、辅助人才评估、领域态势分析等业务实践与研究。

耿哲，管理学硕士，上海交通大学图书馆助理馆员，从事科研竞争情报、科学计量与评价等情报分析服务相关的业务与研究。

前 言

信息素养已提升到国家乃至全球发展战略的高度。信息素养是 21 世纪公民的必备素养,是联合国教科文组织(UNESCO)全民信息计划(Information For All Programme, IFAP)的优先领域。《“十三五”国家信息化规划》指出“支持普通高等学校、军队院校、行业协会、培训机构等开展信息素养培养”“提升国民信息素养”。2009 年 10 月 1 日,美国前总统奥巴马签署代号为 3195-W9-P 的“国家信息素养宣传月”(National Information Awareness Month)议案,指出:“我们还必须学习掌握在任何情况下获取、整理和评价信息的必要的技能,而非仅仅拥有数据”,“查找、找到与辨认信息的能力可以运用于无数的生活决定,无论是财务、医学、教育,还是技术方面。”

提升信息素养最直接、有效的途径是信息检索有关课程的学习。本书定位于信息检索教材,旨在从信息检索的基本概念、原理出发,阐述信息检索的基本理论、方法与技术,介绍国内外重要的综合性与专业性信息检索系统,阐述专利、商标、学位论文、会议论文、科技报告等专类信息检索的技巧以及信息资源选择与评价的方法,分析信息检索与利用中的费用与知识产权问题,介绍个人文献管理软件、自动问答系统与自动翻译系统的使用方法,论述信息检索的重要应用——科技查新与学科前沿分析。

本书先后入选国家“十一五”“十二五”规划教材。其第一版《网络信息检索与利用》(武汉大学出版社,2002 年)自 2002 年出版至 2010 年已重印 10 次,第二版《信息检索》(武汉大学出版社,2010 年)出版后已重印 10 余次,被多所兄弟院系作为本科生教材或硕士生入学考试指定参考书,被 30 余所大学和省级图书馆作为文检课教学、培训、职称考试和业务竞赛的教材或指定用书。

黄如花负责全书的策划和提出编写大纲,胡永生、程银桂参与策划并对大纲提出修改建议。各章节撰写者为:第一章,第二章的第一节与第三节,第五章第一节的 WOS 和 JCR、第二节 SDOL、第五节 OCLC Firstsearch,第

八章,第十章第一节,附录1~3由黄如花编写;第六章第一、二、四节由胡永生在林佳编写的第二版内容基础上修改;第四章,第六章的六至九节由胡永生编写;第九章的第一节与第四节由程银桂编写;第八章由吴丹撰写;第二章第二节,第五章第一节的 InCites 和 ESI、第三节 Scopus、第四节 Scival、第七节 ProQuest 由韩丽编写;第三章和第九章的第二、三节由徐轩编写;第六章第三、五节由陈振英编写;第六章第十节由吴文辉编写;第十章第二节的一、二由杨眉编写,第十章第二节的三由耿哲编写。全书由黄如花审稿和统稿,程银桂对初稿提出修改建议。武汉大学信息管理学院研究生石乐怡、黄雨婷、张俊参与附录和部分章节的资料收集与更新工作。

针对信息检索技术的发展和用户获取信息的变化,也得益于广大读者提出的宝贵意见,与第二版相比,本版有以下几个主要变化:第一,更加突出教材的针对性与应用性。去掉了对本科生难度较大、且发展很快的多媒体信息检索,根据检索手段和信息检索应用的变化,新增了“第八章 移动搜索”和“第十章 第二节 信息检索在学科前沿分析中的应用”;第二,及时反映检索系统的变化、数据库公司之间的合并及图书情报机构的使用现状。鉴于 Dialog 和 CSA 都并入了 ProQuest,去掉了原第二版“第五章 国外综合性信息检索系统”中的“世界上最大的联机检索系统 Dialog”,而在“第五章 国外综合性信息检索系统”新增了 Incites、ESI、Scopus 和 Scival 这四个在国内使用逐渐增多的重要的学术信息检索系统,对第四至六章、“附录1 信息检索有关的网络信息资源”进行了大幅修改;第三,吸取了国家标准《科技查新技术规范》(GB/T 32003-2015)的营养,对科技查新相关术语、程序的相关内容进行了更新,并增加“附录2 查新合同的格式”和“附录3 查新报告的格式”;第四,力图更好地与信息检索实践和文献检索课程的需要相结合,特邀清华大学、上海交通大学、复旦大学、浙江大学、武汉大学、北京师范大学、中南财经政法大学和武汉理工大学等高校从事学科服务或文献检索课程教学的同行参编;第五,更好地服务于信息素养教育领域的教学改革。为了便于文献检索课教学的广大同行备课或者开展翻转课堂教学,本书主编全程独立主讲的“信息检索”MOOC 在中宣部“学习强国”平台“慕课”栏目和教育部主管的爱课程网(<http://www.icourse163.org/course/whu-29001#/info>)上线,提供大量与本教材相关的授课视频、思考题、最新资料等,使之成为真正的立体化教材。

本书在写作过程中借鉴了国内外大量相关研究成果,在此,谨向这些文献的作者致以诚挚的谢意!本书多个版本的出版均得到武汉大学出版社的支持,责任编辑詹蜜为本书的出版付出了辛勤劳动,提出了许多宝贵的修改意

见，在此谨表谢忱！

信息检索技术和检索工具日新月异，加之编著者学识、水平有限，书中疏漏和不妥乃至错误之处在所难免，敬祈专家学者和读者批评指正，以便今后的修订和补充。

黄如花

2019年8月于武昌珞珈山

目 录

第一章 信息检索概述.....	1
第一节 信息检索的基本概念.....	1
第二节 信息检索的历史.....	4
第三节 信息检索的模型.....	9
第四节 信息检索系统的结构与评价	12
第二章 信息检索的方法与技术	17
第一节 网络信息检索的基本方法	17
第二节 信息检索的主要技术	24
第三节 信息检索的技巧	29
第三章 搜索引擎	35
第一节 搜索引擎概述	35
第二节 综合性搜索引擎选介	41
第三节 中外文学学术搜索引擎选介	44
第四章 国内重要的综合性信息检索系统	46
第一节 中国知网(CNKI)	46
第二节 维普资讯网(VIP)	54
第三节 国家科技图书文献中心(NSTL)	60
第四节 中国高等教育文献保障系统(CALIS)	65
第五节 中国科学院国家科学图书馆	68
第五章 国外重要的综合性信息检索系统	73
第一节 学术资源整合平台——Web of Science	73

第二节	ScienceDirect Online (SDOL)	81
第三节	Scopus	85
第四节	Scival	93
第五节	FirstSearch	97
第六节	Gale 数据库检索系统	99
第七节	ProQuest 系列数据库平台	102
第六章	国内外专业性书目信息检索系统	115
第一节	SciFinder	115
第二节	剑桥科学文摘(CSA)	120
第三节	生物学文摘(BA)	124
第四节	工程索引(Ei Compendex)	126
第五节	美国医学文摘(Medline)	134
第六节	荷兰医学文摘	137
第七节	教育资源信息中心(ERIC)	142
第八节	公共事务信息数据库(PAIS International)	145
第九节	图书情报学专业数据库——LISA 和 LISTA	149
第十节	法学专业数据库——Lexis Advance	153
第七章	专类信息的检索	160
第一节	专利信息的检索	160
第二节	商标信息的检索	177
第三节	学位论文信息的检索	183
第四节	会议论文信息的检索	193
第五节	科技报告的检索	197
第八章	移动搜索	205
第一节	移动搜索背景	205
第二节	移动搜索类型与特点	210
第三节	移动搜索工具与评价	214
第四节	移动搜索发展趋势	217
第九章	网络信息检索与利用中的有关问题	223
第一节	网络信息的选择与评价	223

第二节 网络信息检索与利用中的费用问题.....	235
第三节 网络信息检索与利用中的知识产权保护.....	238
第四节 网络信息检索与利用的重要工具.....	240
第十章 信息检索的主要应用.....	250
第一节 信息检索在科技查新中的应用.....	250
第二节 信息检索在学科前沿分析中的应用.....	267
附录1 信息检索有关的免费网络资源	285
附录2 查新合同的格式	295
附录3 查新报告的格式	299
主要参考文献	305

图 目 录

图 1-1	信息检索的基本原理	3
图 1-2	计算机检索系统的逻辑结构	15
图 2-1	文氏图：逻辑“与”	18
图 2-2	文氏图：逻辑“或”	18
图 2-3	文氏图：逻辑“非”	18
图 2-4	信息检索的流程	29
图 2-5	积木型检索式构造	32
图 4-1	中国知网的一框式检索界面	49
图 4-2	中国知网的高级检索界面	49
图 4-3	CNKI 的专业检索界面	51
图 4-4	CNKI 的开放获取学术关注度示例	52
图 4-5	CNKI 开放获取 2016 年热点文献	54
图 4-6	CNKI 开放获取媒体关注度示例	54
图 4-7	CNKI 开放获取学术传播度示例	54
图 4-8	维普资讯的一框式检索实例	55
图 4-9	维普资讯的高级检索界面	56
图 4-10	维普资讯的高级检索结果界面	57
图 4-11	维普资讯的期刊检索模式	58
图 4-12	维普资讯的文献引证追踪检索实例	58
图 4-13	维普资讯的被引用论文实例	59
图 4-14	维普资讯的引证文献实例	59
图 4-15	维普资讯的引用追踪实例	60
图 4-16	NSTL 的文献检索界面	62
图 4-17	NSTL 的智能检索界面	63
图 4-18	NSTL 的国际科学引文数据库	64

图 4-19	CALIS 的 e 读学术搜索检索界面	67
图 4-20	CALIS 的 e 读学术搜索检索结果	67
图 4-21	CSDL 资源集成发现系统检索界面	69
图 4-22	CSDL 资源集成发现系统检索实例	70
图 4-23	CSDL 资源集成发现系统跨库高级检索界面	70
图 4-24	CSDL 资源集成发现系统检索结果界面	71
图 5-1	Scopus 文献检索界面	87
图 5-2	Scopus 的作者搜索页面	88
图 5-3	Scopus 归属机构搜索页面	89
图 5-4	Scopus 高级搜索页面	89
图 5-5	Scopus 的来源出版物页面	90
图 5-6	Scopus 的导出文献设置	91
图 5-7	Scival 首页	94
图 5-8	Scopus 的 Topics 轮状呈现图	96
图 5-9	ProQuest 平台选择数据库页面	107
图 5-10	ProQuest 基本检索界面	109
图 5-11	ProQuest 高级检索界面	110
图 5-12	ProQuest 词库检索实例	111
图 5-13	ProQuest 命令行检索界面	112
图 6-1	CSA 基本检索界面与实例	121
图 6-2	CSA 高级检索的界面与实例	122
图 6-3	CSA 命令行检索方式的界面与实例	122
图 6-4	CSA 检索结果界面	124
图 6-5	Ei 快速检索界面	127
图 6-6	Ei 的专家检索界面	130
图 6-7	Ei 的叙词检索界面	130
图 6-8	Ei Compendex 检索结果页面	132
图 6-9	Ei 的基金组织分布图	133
图 6-10	Ei 检索结果下载设置页面	133
图 6-11	Embase 的搜索界面	138
图 6-12	Embase 的高级检索界面	138
图 6-13	Embase 快速检索实例	139
图 6-14	Emtree 主题词表检索	139
图 6-15	Embase 主题词表浏览界面	140

图 6-16	Embase 的期刊浏览界面	141
图 6-17	Embase 的检索结果界面	141
图 6-18	Embase 的 My Tools 选项	142
图 6-19	ERIC 检索界面	143
图 6-20	ERIC 检索实例	144
图 6-21	ERIC 叙词表检索实例	144
图 6-22	ERIC 叙词表中的上位词、下位词和相关词	145
图 6-23	PAIS 快速检索界面与实例	146
图 6-24	PAIS 高级检索界面与实例	147
图 6-25	PAIS 检索结果界面	148
图 6-26	EBSCOhost 的基本检索界面	151
图 6-27	EBSCOhost 的高级检索界面	151
图 6-28	EBSCOhost 的辅助检索功能界面	152
图 6-29	LISTA 的词表浏览功能	152
图 6-30	EBSCOhost 的检索结果页面	153
图 6-31	Lexis Advance 的检索页面	155
图 6-32	Lexis Advance 的判例文阅读页面	156
图 6-33	Lexis 的 Shepard's citation 服务	157
图 10-1	查新质量评价指标体系	266
图 10-2	前沿研究态势检索与分析流程	268
图 10-3	特定群体的人才检索与分析流程	276
图 10-4	机构科研竞争力检索与分析流程	279

表 目 录

表 1-1	信息检索的模型	10
表 2-1	常用字段名及其代码	21
表 3-1	重要的中英文综合性搜索引擎	42
表 4-1	CNKI 专业检索运算符使用说明	50
表 4-2	维普资讯网的主要资源	54
表 4-3	维普资讯的期刊检索模式说明	57
表 5-1	SDOL 高级检索的可检字段及其含义	83
表 5-2	Scopus 四大类 27 个一级学科类目	86
表 5-3	OCLC FirstSearch 的基本组数据库	98
表 6-1	SciFinder 检索模式和检索途径	116
表 6-2	CSA 检索运算符	123
表 6-3	Ei 快速检索区字段说明	128
表 6-4	Ei 检索限制条件说明: 文献类型	129
表 6-5	Ei 检索限制条件说明: 文献处理类说明	129
表 6-6	Ei 位置算符	131
表 6-7	PAIS 检索运算符	147
表 6-8	LISTA 检索运算符	149
表 7-1	Clarivate Analytics 和 ProQuest Dialog Solution 提供服务的 专利数据库	170
表 7-2	提供专利信息或专利检索的国内外知识产权管理机构网站	172
表 7-3	可免费获取国外学位论文的重要站点(按国家或地区 集中排列)	191
表 7-4	提供免费科技报告的站点	202
表 8-1	移动搜索与桌面搜索的比较	213
表 9-1	免费在线翻译网站(按功能多少排列)	247

第一章 信息检索概述

第一节 信息检索的基本概念

一、信息检索的含义

信息检索作为一种实践活动由来已久。但作为一个比较规范、正式的学术术语,信息检索(Information Retrieval,简称IR)这个术语1950年由美国信息科学的先锋Calvin Northrup Mooers(1919—1994)首先提出,这也是他1978年获得美国信息科学协会荣誉奖(American Society for Information Science's Award of Merit)的原因之一^①。

广义的信息检索是指将信息按一定的方式组织和存储起来,并根据信息用户的需要找出有关信息的过程。所以,它的全称又叫信息存储与检索(information storage and retrieval),即包括信息的“存”与“取”两个环节。广义信息检索的其他表述有:信息检索是对信息项(information items)进行表示(representation)、存储(storage)、组织(organization)和存取(access)^②。

狭义的信息检索则仅指该过程的后半部分,即从信息集合中找出所需信息的过程,相当于“信息查询”或“信息查找”(information search)。“信息检索”的含义很广,但作为一个学术研究领域,可界定为:信息检索是从文档集合(通常存储在计算机中)查找满足某种信息需求的具有非结构化性质(通常

^① Calvin Northrup Mooers[EB/OL]. [2019-07-25]. <https://www.asist.org/pioneers/calvin-northrup-mooers/>.

^② Ricardo Barza Yates, Berthier Ribeiro Neto, et al. Modern Information Retrieval[M]. Boston: Addison Wesley Longman Publishing Co. Inc., 1999: 1.

指文本)的资料(通常是文献)。”^①可见,这也是从狭义的角度界定的。在通常情况下,人们讲“信息检索”是从狭义的角度而言的。

信息检索与文献检索的主要区别在于:文献检索是以获取文献信息为目的的检索,信息检索则收集、组织、存储一定范畴的信息,并可供用户按需要查询文献中的信息或知识单元,比文献检索更深入。

二、信息检索的种类

根据检索手段的不同,信息检索可分为手工检索、光盘检索、联机检索和网络检索。网络检索是当前信息检索的主流和发展方向,因而本书以网络检索为主。

根据检索对象形式的不同,信息检索又可分为:

文献型信息检索(document retrieval):是以文献(包括题录、文摘和全文)为检索对象的检索。凡是查找某一主题、时代、地区、著者、文种的有关文献,以及这些文献的出处和收藏处所等,都属于文献型信息检索的范畴。完成文献型信息检索主要借助于各种书目型数据库。

数值型信息检索(data retrieval):是以数值或数据为对象的一种检索,包括文献中的某一数据、公式、图表,以及某一物质的化学分子式等,数据检索分为数值型与非数值型。完成数值型信息检索主要借助于各种数值数据库和统计数据库。

事实型信息检索(fact retrieval):是以某一客观事实为检索对象,查找某一事物发生的时间、地点及过程的检索,其检索结果主要是客观事实或为说明事实而提供的相关资料。完成事实型信息检索主要借助于各种指南数据库和全文数据库。

关于从检索对象的角度对信息检索类型的划分,出现了一种新的三分方法,即文本检索、数值检索、音频与视频检索。相对于早期对信息检索概念的细分方法,新的三分法比较全面地反映了信息检索概念的基本内涵和最新发展^②。

① Christopher D. Manning, Prabhakar Raghavan, Hinrich Schütze. An Introduction to Information Retrieval(Online edition)[M]. Cambridge: Cambridge University Press, 2009: 1.

② 赵丹群编著. 现代信息检索:原理、方法与技术[M]. 北京:北京大学出版社, 2008: 2.