



面向 21 世纪 课程教材
Textbook Series for 21st Century

信息组织

(第二版)

戴维民 主编



高等教育出版社
HIGHER EDUCATION PRESS

附赠学习卡

面向 21 世纪课程教材
Textbook Series for 21st Century

信息组织

(第二版)

戴维民 主编



高等教育出版社
HIGHER EDUCATION PRESS

图书在版编目(CIP)数据

信息组织/戴维民主编. —2版. —北京:高等教育出版社, 2009. 5

ISBN 978-7-04-026655-9

I. 信… II. 戴… III. 信息管理-高等学校-教材
IV. G203

中国版本图书馆 CIP 数据核字(2009)第 056828 号

策划编辑	罗雪群	责任编辑	王友富	封面设计	杨立新
责任绘图	尹文军	版式设计	王莹	责任校对	胡晓琪
责任印制	尤静				

出版发行 高等教育出版社
社 址 北京市西城区德外大街 4 号
邮政编码 100120
总 机 010-58581000

经 销 蓝色畅想图书发行有限公司
印 刷 化学工业出版社印刷厂

购书热线 010-58581118
免费咨询 800-810-0598
网 址 <http://www.hep.edu.cn>
<http://www.hep.com.cn>
网上订购 <http://www.landraco.com>
<http://www.landraco.com.cn>
畅想教育 <http://www.widedu.com>

开 本 787×960 1/16
印 张 27.5
字 数 510 000

版 次 2004 年 7 月第 1 版
2009 年 5 月第 2 版
印 次 2009 年 5 月第 1 次印刷
定 价 35.40 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 26655-00

作者简介

戴维民 南京政治学院上海分院院长,教授,文学博士,博士生导师,兼任中国索引学会常务理事,中国图书馆学会常务理事,国际知识组织学会会员,教育部高等学校档案学学科教学指导委员会副主任委员,上海市图书馆学会副理事长、《中国图书馆分类法》编委会委员等。享受国务院政府特殊津贴。

主要学术领域是信息传播、信息组织、信息资源管理、军队政治工作信息化。完成国家级、军队级等各类课题十多项,出版著作 20 余部,发表论文 180 多篇。主要著作有《情报检索语言综论》、《网络信息优化传播导论》、《语义网信息组织技术与方法》等。主要论文有《论我国情报检索语言研究道路》等。

曹树金 中山大学资讯管理系主任、教授、博士生导师,是教育部“新世纪优秀人才支持计划”入选人。兼任中国索引学会副理事长、中国图书馆学会理事、中国图书馆学会标引与编目专业委员会委员、广东图书馆学会副理事长及学术委员会主任委员、广东科技情报学会常务理事、广东社科情报学会理事、《中国图书馆分类法》编委。主要研究领域有情报检索语言与信息组织、用户信息行为、网络信息管理、市场营销、电子商务。主持和参与国家级、省部级科研课题十多项,发表论文 80 多篇。主要著作有《信息组织的分类法与主题法》。主要论文有《DDC、UDC 网络化对〈中图法〉发展的启示》、《用户网络信息查询需求研究》等。

陈树年 华东理工大学研究馆员,中国图书馆学会学术委员、标引与编目专业委员会副主任,全国信息与文献标准化委员会委员,国际知识组织学会会员,《中国图书馆分类法》副主编,信息产业部网络信息组织标准工作组成员。曾主持修订国家标准《文献分类标引规则》、《文献主题标引规则》,主持制订电子行业标准《网络信息分类系统》。出版著作 20 多部(含参编),发表论文 60 多篇。主要著作有《中国分类主题词表使用手册》、《中图法理论与使用》、《大学文献信息检索教程》等。主要论文有《搜索引擎与网络信息资源的分类组织》、《面向 21 世纪〈中图法〉及其标引技术改造的研究》等。

包冬梅 南京政治学院上海分院讲师,管理学硕士。参与多项国家级、军队级课题的研究,参编著作多部,发表论文近20余篇。主要论文有《Web资源自然语言检索后控制优化技术》、《传统图书情报技术在网络信息组织和检索中的应用和发展》等。

张建忠 南京政治学院上海分院讲师,文学硕士。主要研究方向有情报检索语言、信息资源管理。主要著作有《知识的组织》、《知识的揭示》。主要论文有《比较图书馆学研究综述》、《论“著作权法”对图书馆馆藏建设的影响与对策》等。

王珏 华东师范大学信息学系讲师,博士。研究方向主要有信息资源组织与利用、日本图书馆学。主要著作有《日语能力测试基本词汇例解》、《基于Web的信息服务 OCLC FirstSearch 服务分析》(合著)。主要论文有《日本科学文献数据库及其服务的发展》、《论日本的图书馆情报学专业资格制度》等。

罗春荣 中山大学图书馆研究馆员、副馆长、中山大学硕士生导师。中国高等教学文献保障体系华南地区中心管理委员会委员,广东图书馆学会信息技术委员会副主任委员,广东省科学技术情报学会常务理事。主要研究领域有电子馆藏发展与管理、数字图书馆建设、信息检索、信息公开。主持参与多项省部级、国家级课题研究,发表论文30多篇。主要著作有《文献信息检索与利用》(合著)。主要论文有《电子馆藏及其发展政策研究》、《电子馆藏评估:内容与方法》、《主题检索语言的术语学基础》等。

甘英 山东大学教授,曾任山东省图书馆副馆长。主要著作有《为学之道》、《信息资源与信息收集》(合著)等。主要论文有《山东省图书馆分类法述略》、《关于信息收集及信息收集学的探讨》、《竞争情报收集中人的因素》等。

邱君瑞 第二军医大学副研究馆员,管理学硕士。主要论文有:《主题词表在网络检索系统中的应用调查分析》、《DDC的网络适应性改造对中医法未来发展的启示》、《论一体化医学语言系统超级叙词表的概念表达》。

罗昊 武警上海政治学院讲师,管理学博士。主要论文有《语义网信息组织机制论纲》、《聚焦信息组织研究的前沿——面向网络信息组织的本体开发》、《网络本体语言构成的理论模型研究》。

第一版作者名单

主 编

戴维民 南京政治学院上海分院

编 委

曹树金 中山大学

陈树年 华东理工大学

包冬梅 南京政治学院上海分院

张建忠 南京政治学院上海分院

王 珏 华东师范大学

罗春荣 中山大学图书馆

甘 英 山东大学

邱君瑞 第二军医大学

第二版修订说明

本教材第一版出版至今已经5年了,记得5年前的春节,为了赶在规定的时间内交稿,本书的作者们忙活了一个春节。大家的努力和付出是有成果的,因为本教材出版以来被多所院系采用,并被列入考研必读参考书,受到了教授和学习本课程师生的好评,我们编写人员深感欣慰和自豪。

过去的5年,信息组织的理论和技术有了新的发展,信息组织作为一个分支学科被广泛认可,并且成为一个比较活跃的学术领域,出版了大批的著作、教材和论文。与此同时,信息组织实践也有了新的进步,信息组织技术和方法不断发展,特别是网络信息组织技术、自然语言信息检索技术发展迅速,为信息组织课程的教学提供了许多新的内容。因此,本教材需要与时俱进地加以修订,以与理论和技术的发展保持同步。这次的修订除了对第一版少量不够准确或因时代发展而过时的内容进行校订外,及时补充了信息组织领域的新理论、新知识、新技术和新方法,对网络信息组织、自然语言信息检索方法等方面的内容作了比较大的增补,力图教授和学习信息组织的人们提供最新的理论与方法。

第一版出版后,我们制作了与之配套的网络课程,以适应网络时代的学习需要,为教师备课和学生深度拓展学习提供了更大的方便,受到了师生们的欢迎,我们决定,在本教材第二版出版后同步修订网络课程,及时更新网络资源。

本次修订仍按第一版的撰写分工基本上由原作者完成,罗昊参加了第八章的修订,包冬梅对附录进行了修订更新,并协助主编做了大量的协调工作和统稿工作。高等教育出版社非常重视本教材修订工作,给予及时的指导和支持,罗雪群编辑春节过后一上班就专程到上海组织审稿统稿。在此谨向全体作者和高等教育出版社的同志表示衷心的感谢。

这次修订又逢牛年春节,这个春节我们也是与它共同度过的,相信我们的努力和付出都是有意义的。

戴维民

二〇〇九年春节于上海

第一版前言

在我国近现代图书馆学史和图书馆学教育史上,一些非常著名的图书馆学家,如刘国钧、杜定友、皮高品等,他们既编制了一些有影响的分类法,同时也曾在大学课堂上讲授图书分类课程。图书分类既是图书馆重要的技术工作,也是图书馆学的重要教学内容。正因为如此,图书分类学就诞生了。后来,图书馆馆藏不仅有图书,还有各种其他文献,因而有了文献分类学。

20世纪70年代,我国主题法有了初步发展,逐步形成了分类法和主题法两大方法体系并存的格局,在大学教学中分设了专门课程。但在20世纪80年代初,张琪玉教授开始了对分类法、主题法等各种信息组织方法的综合研究,规范了名词术语,为图书馆学情报学开拓了一个全新的研究领域——情报检索语言,并成功地开设了相应的课程。后在此基础上又系统创立和发展了情报语言学的学科体系。这是一门理论与实践紧密结合、中外兼容的图书馆学情报学分支学科。20多年的创建发展和完善,不仅使之走向成熟,某种程度上也创立了一个学派。

进入21世纪,互联网技术更加普及和发展,数字图书馆技术、网络信息检索技术给我们带来了许多全新的课题,或许是追赶时尚的需要,广泛采用了“信息组织”这一名称,并进而成为大学的课程。这在信息管理与信息系统专业已经先行一步,教育部图书馆学学科教学指导委员会审时度势,决定组织编写图书馆学专业核心课程系列教材,《信息组织》忝列其中,这便是这本书的来历。课程名称从图书分类学到分类法与主题法、情报检索语言,再发展为信息组织,标志着教学内容的变化,而且与世界主要国家同类专业课程建设保持了同步发展。我访问了一些美国大学的图书馆学信息学研究生院的网站,从他们的课程计划表中,可以看出几乎各大学都有“信息组织”或类似的课程,例如亚拉巴马大学图书馆与信息学研究生院、亚利桑那大学信息资源与图书馆学研究生院、肯特州立大学图书馆与信息学研究生院开设了“信息组织”(Organization of Information)课程,圣何塞州立大学图书馆与信息学研究生院开设了“信息组织与管理”(Information Organizations and Management)、“高级编目与信息组织”(Advanced Cataloging and Organization of Information)课程,肯塔基大学图书馆与信息学研究生院开设了“知识组织(一)”(Organization Knowledge I)、“知识组织(二)”(Organization Knowledge II)课程,罗杰斯大学传播、信息与图书馆学研究生院开设了“信息组织法”(Organizing Information)、“信息检索的知识表示”

(Knowledge Representation for Information Retrieval)课程,匹兹堡大学信息学研究生院开设了“信息组织法”课程等。由此,我相信这是世界范围内图书馆学情报学大学生和研究生课程的共同发展趋势。

我非常荣幸地担任了这部教材的主编,并得到了全体编写者的大力支持和帮助。本教材由我提出框架大纲,然后集思广益地完成了详细的编写大纲,据此分工写作。第一、二章由南京政治学院上海分院戴维民编写,第三章由华东理工大学陈树年编写,第四章由中山大学罗春荣、曹树金编写,第五章由南京政治学院上海分院张建忠编写,第六章由中山大学曹树金、罗春荣编写,第七章由南京政治学院上海分院包冬梅编写,第八章由戴维民、包冬梅和第二军医大学邱君瑞编写,第九章由华东师范大学王珏编写,第十章由山东大学甘英编写,附录一由邱君瑞编写,附录二、三由戴维民编写,最后由戴维民统稿定稿。

为了赶在规定的时间内交稿,各位编写者在2004年春节期間都放弃了休息,共同努力完成了编写任务。但由于多人分头编写,而且时间非常紧张,我未能有足够的时间去做统一风格、认真雕琢的统稿工作,因此,本教材一定有许多不足之处,希望得到大家的批评指正,也希望在修订再版时作进一步完善。

戴维民

2004年3月于上海

内 容 提 要

本书是教育部“面向 21 世纪课程教材”，是由教育部高等学校图书馆学学科教学指导委员会组织编写的图书馆学专业核心课程系列教材之一。

本书根据现代信息资源的特点以及信息检索的要求，对信息组织的方法进行了系统全面的描述，内容包括：现代信息组织的背景和要求、信息组织的基本原理、信息组织分类法、信息组织主题法、信息组织集成法、信息内容分析与标引、信息组织中的自然语言应用、网络信息组织、数字图书馆信息组织、信息组织的历史发展与未来趋向。第二版修订时根据近几年信息组织理论与技术的发展又进行了及时更新和校订，特别是对网络信息组织、自然语言信息检索方法等发展快速的领域进行了较大篇幅的增补，比较全面地反映了信息组织最新的理论与方法。本书的特点是：理论与实践紧密结合，将传统文献组织与现代信息组织融为一体，既面向现实也展望未来。书后附有中外重要信息组织工具简介、“信息组织”课程推荐书目和信息组织相关的网络资源指南，可以为教学提供更多的参考。与第一版书配套的网络课程为教师备课和学生的深入学习提供了极大的方便，深受教授和学习信息组织课程的师生的欢迎，此次教材修订也将同步对网络课程进行修订，并努力保持适度的动态更新。

本书可作为高等学校图书馆学、信息管理及相关专业教学用书，也可作为各类图书馆和信息机构岗位培训教材和工作参考书。



目 录

第一章 绪论	1
第一节 信息组织环境	1
一、信息空间的信息环境	1
二、信息检索需求的多样化	2
三、信息检索机制的变化	3
四、信息检索工具的变化	4
五、信息检索技术的进步	5
第二节 信息组织方法体系	6
一、信息揭示与信息组织	6
二、信息组织方法	7
三、信息组织手段	8
第三节 知识组织的新时代	9
一、知识经济与知识管理的新时代	9
二、从信息组织走向知识组织	11
第二章 信息组织的基本原理	14
第一节 语言学原理	14
一、信息组织的语言工具	14
二、检索语言的类型	15
三、语言学原理在信息组织中的应用	19
四、检索语言的语法	21
第二节 系统论原理	28
一、系统与系统论	28
二、信息检索系统的发展	29
三、信息检索系统的结构	31
四、信息检索系统评价	33
第三节 知识分类原理	35
一、检索语言的知识体系的来源	35
二、哲学家知识分类体系	36
三、大百科全书知识分类体系	38
四、大学教学知识分类体系	39

五、科学学的知识分类	39
第四节 概念逻辑原理	40
一、概念种类、内涵与外延	40
二、概念之间的关系	42
三、概念逻辑方法	43
四、概念关系的表达	44
第三章 信息组织分类法	48
第一节 分类法的原理	48
一、信息分类检索的特点与需求	48
二、信息分类法的类型	51
三、信息分类法的结构原理	59
四、分类法结构与功能的关系	66
第二节 分类法的编制	68
一、分类法编制的基本程序	68
二、分类体系的构建	70
三、类目的划分与排列	71
四、类目关系显示及交叉关系的处理	81
五、类目注释	84
六、分类法的修订	85
第三节 分类法的标记系统	87
一、分类法标记的功能与要求	87
二、分类法标记的类型	89
三、分类法标记系统的构建	94
第四节 《中国图书馆分类法》概述	100
一、《中图法》的历史发展	100
二、《中图法》的结构原理	102
三、《中图法》的电子版	108
第四章 信息组织主题法	114
第一节 主题法的原理与功能	114
一、主题检索的特点与需求	114
二、主题法的类型	115
三、主题法的原理	119
四、主题法的功能	121
第二节 叙词法的词汇控制	122
一、词汇选择	122

二、词形控制与词义控制	124
三、词间关系控制	128
第三节 叙词表的结构与编制	133
一、叙词表的宏观结构	133
二、叙词表的微观结构	136
三、叙词表的编制与维护	143
第五章 信息组织集成法	149
第一节 信息组织方法的融合	149
一、信息组织工具的结构与功能	149
二、信息组织方法融合的形式	151
第二节 信息组织工具兼容互换	155
一、信息组织工具的主动式兼容	156
二、信息组织工具的被动式兼容	157
第三节 分类主题一体化	160
一、分类主题一体化发展趋向	160
二、分类主题一体化词表的结构原理	162
三、分类主题一体化词表的类型	163
四、分类主题一体化方法的应用	165
第四节 《中国分类主题词表》的结构与使用	167
一、《中国分类主题词表》的编制概况	167
二、《中国分类主题词表》体系结构	169
三、《中国分类主题词表》电子版	170
四、分类号—主题词对应表	176
五、主题词—分类号对应表	179
六、《中国分类主题词表》在分类标引中的应用	181
七、《中国分类主题词表》在主题标引中的应用	184
第六章 信息内容分析与标引	187
第一节 信息主题与主题分析	187
一、信息主题的含义、结构与类型	187
二、主题分析的水平和要求	192
第二节 信息标引的种类、方式和程序	195
一、信息分类标引和主题标引	195
二、各种信息标引方式	196
三、标引工作程序	199
第三节 分类标引的方法与规则	201

一、辨类的方法	201
二、分类标引的基本规则	204
三、分类标引的一般规则	205
四、各学科信息的分类标引规则	215
五、确定分类法使用本与图书改编	215
六、同类书区分	217
第四节 主题标引的方法与规则	221
一、主题概念分解与查表选词的方法	221
二、选择标引词的一般规则	227
三、各类型主题与各类型文献的主题标引规则	230
第七章 信息组织中的自然语言应用	240
第一节 自然语言在信息组织中的应用概述	240
一、自然语言的演化	240
二、自然语言区别于受控语言的特点	242
三、自然语言处理及其在信息组织和检索中的应用	243
第二节 自然语言在信息组织中的应用	244
一、自动标引的实现基础——自动分词	244
二、自然语言标引	250
三、自动分类	255
第三节 自然语言检索系统与自然语言检索	258
一、自然语言检索系统概述	258
二、自然语言检索	259
三、全文检索	260
四、搜索引擎的自然语言检索	263
五、自然语言检索系统的优点及存在的不足	267
第四节 后控制检索	268
一、后控制机制概述	268
二、国内外后控词表研究及其应用现状	271
三、网络检索系统中的后控制技术	275
第八章 网络信息组织	281
第一节 网络信息类型与特点	281
一、网络信息及其类型	281
二、网络信息的环境特点和组织的难点	283
三、网络信息组织的目标	284
第二节 网络信息的分类组织	285

一、传统分类法应用于网络信息组织	285
二、网络信息分类法与分类模式	287
三、网络信息自动分类	291
第三节 网络信息的主题组织	295
一、基于网络的叙词表的发展	295
二、叙词表在网络多媒体信息组织中的应用	297
三、主题法在网关中的应用	298
第四节 基于本体的信息组织	299
一、语义网信息组织新模式	299
二、本体基本原理	301
三、网络本体描述语言	304
四、本体构造	307
五、本体标注	311
第五节 网络信息组织方式	312
一、文件方式	312
二、数据库方式	313
三、主题树方式	314
四、搜索引擎方式	316
五、Web2.0 信息自组织方式	319
第六节 网络信息重组与知识挖掘	321
一、网络信息重组与导航	321
二、网络知识挖掘	327
第九章 数字图书馆信息组织	331
第一节 数字图书馆信息组织概述	331
一、数字图书馆的发展	331
二、数字图书馆的含义	333
三、数字图书馆信息资源特点	334
四、数字图书馆信息组织的特点	335
第二节 数字图书馆信息组织的基本原理与方法	336
一、数字图书馆系统的基本结构	336
二、数字信息资源的信息结构	337
三、数字图书馆用户获取信息的模式	338
四、信息资源组织的基本原则	339
五、数字图书馆信息资源组织步骤	340
第三节 数字信息资源的描述	341

一、元数据的定义、类型	341
二、DC	342
三、MODS	345
四、XML 与 RDF	349
第四节 数字信息资源的标引与整合	357
一、数字信息资源标引与整合的含义	357
二、数字信息资源的分类标引与主题标引	358
三、非文本信息资源的标引与整合	359
四、异构数字资源库的整合	360
第十章 信息组织的历史发展与未来趋向	363
第一节 古代信息组织	363
一、信息组织产生的基础	363
二、古代的信息组织	364
第二节 近代信息组织	367
一、近代国外信息组织	367
二、我国近代的信息组织	369
三、近代信息组织的特点	370
第三节 现代信息组织	371
一、现代信息组织方法的发展	371
二、网络信息资源组织的探讨与实践	375
三、现代信息组织的特点	378
第四节 信息组织的发展趋势	379
一、信息组织环境的变化	379
二、以用户需求为导向的信息组织	380
三、信息内容揭示深入化	381
四、信息组织方法集成化	381
五、信息组织技术智能化	382
六、信息组织系统兼容化与标准化	383
七、信息组织工具易用化	383
八、信息组织理论研究的拓展	384
主要参考文献	386
附录一 中外重要信息组织工具简介	390
附录二 “信息组织”课程推荐书目	411
附录三 信息组织相关的网络资源指南	416

第一章 绪 论

信息组织是通过一定的工具和技术将无序的信息组织成一个有序系统的方法。信息组织活动是随着信息数量规模的增长和信息检索需求的变化而不断产生和发展的。从几千年前对图书文献的分类,到近现代信息组织工具——检索语言的兴旺发展,再到今天网络信息组织技术的进步,表明信息组织是我们有效利用信息的重要方法。

第一节 信息组织环境

一、信息空间的信息环境

今天,我们面临的是信息爆炸和信息污染的复杂的信息环境。

首先,过去的信息缺乏已变为今天的信息过剩或信息泛滥。世界信息量呈爆炸式增长,1995年的全球总信息量是1985年的2 400倍,现在一天的信息总量相当于1985年全年信息总量的6.5倍。^①这种现象被称为信息爆炸或信息超载。

信息超载(information overload)就是指系统或个人所接受的信息超过其自身的处理能力或信息未能有效利用的状况。信息超载就是人们拥有过多的没有消化或不可能消化的信息。信息超载的结果导致信息生态环境的变化,信息接受效率降低,信息浪费。

信息资源的奇特性在于,不仅信息可以被无损耗地消费,而且信息的消费过程可能同时就是信息的生产过程,信息所包含的知识或感受在消费者那里催生出更多的知识和感受,消费它的人越多,它所包含的资源总量就越大。

从20世纪60年代开始,美国科学家普赖斯(D. Price)对科技文献增长规律进行了研究,提出了“指数增长规律”,即科学文献增长与时间成指数函数关系。20世纪90年代以来,互联网使信息的消费者数量增加到最大限度,因为互联网是一种传播与反馈同时进行的交互性媒介,所以以太网的发明者梅特卡夫(Bob Metcalfo)揭示了网络价值的秘密,即网络的价值与网络的使用者的平方成正比,这就是著名的“梅特卡夫定律”。我们据此可以延伸理解为,随着上网人数的

^① 陈挺,赫兢. 信息安全与国家安全[J]. 首都信息化,2000(3).