



普通高等教育“十一五”国家级规划教材

21世纪 信息资源管理系列教材

信息组织与信息构建

主编/周晓英

副主编/马文峰

齐虹

中国人民大学出版社

普通高等教育“十一五”国家级规划教材
21世纪信息资源管理系列教材

信息组织与信息构建

主 编 周晓英
副主编 马文峰 齐 虹

中国人民大学出版社
· 北京 ·

图书在版编目 (CIP) 数据

信息组织与信息构建/主编周晓英

北京: 中国人民大学出版社, 2010

(普通高等教育“十一五”国家级规划教材; 21 世纪信息资源管理系列教材)

ISBN 978-7-300-12766-8

- I. ①信…
- II. ①周…
- III. ①信息管理-高等学校-教材
- IV. ①G203

中国版本图书馆 CIP 数据核字 (2010) 第 187032 号

普通高等教育“十一五”国家级规划教材

21 世纪信息资源管理系列教材

信息组织与信息构建

主 编 周晓英

副主编 马文峰 齐 虹

Xinxi Zuzhi yu Xinxi Goujian

出版发行	中国人民大学出版社		
社 址	北京中关村大街 31 号	邮政编码	100080
电 话	010-62511242 (总编室)		010-62511398 (质管部)
	010-82501766 (邮购部)		010-62514148 (门市部)
	010-62515195 (发行公司)		010-62515275 (盗版举报)
网 址	http://www.crup.com.cn		
	http://www.ttrnet.com (人大教研网)		
经 销	新华书店		
印 刷	北京密兴印刷有限公司		
规 格	170mm×228mm 16 开本	版 次	2011 年 1 月第 1 版
印 张	18.5	印 次	2011 年 1 月第 1 次印刷
字 数	321 000	定 价	29.80 元

版权所有 侵权必究 印装差错 负责调换

作者简介

周晓英，女，1963年出生于湖北省武汉市，北京大学管理学博士，现为中国人民大学信息资源管理学院教授、博士生导师，教育部2007新世纪优秀人才。主要研究方向是信息构建、信息组织、信息资源管理、政府网站建设、政府信息资源管理、情报学理论与方法、信息获取和信息分析。

主要社会兼职：中国社科情报学会副秘书长、常务理事、学术委员会委员；中国科技信息学会理论方法与教育委员会委员；北京市信息化专家委员会委员；北京社会科学信息学会理事；国家信息资源管理北京研究基地客座研究员。复印报刊资料《图书馆学情报学》杂志执行主编、《情报科学》、《情报资料工作》编委会委员、中科院电子政务理事会理事、《电子政务》杂志学术委员会委员。

获奖：至2009年共两次独立获得省部级优秀科研成果奖（二等奖和三等奖各一个）。

成果：在国内外发表学术论文80余篇，主编和独撰学术专著4部，出版教材4部，参与编写书籍多部。

齐虹，女，1965年出生，中国人民大学管理学博士，现为中国人民大学信息资源管理学院副教授、硕士生导师。主要研究方向是信息组织、信息服务与市场、信息用户。

科研成果：主要著作有《政府信息资源管理》（合著）。发表论文主要有《信息服务原理研究》、《信息服务的成本收益效应解析——兼议信息服务机制运行的经济学理由》、《信息选择——信息服务功能拓展的哲学基础》、《论政府信息服务中的正反馈机制》等20余篇。参与编写出版教材一部。

马文峰，女，1953年生。中国人民大学图书馆研究馆员，中国人民大学信息资源管理学院兼职教授，硕士生导师。出版著作（独/参编）8部，发表学术论文70余篇。

内容简介

信息组织与信息构建是网络环境下信息管理的核心内容，是具有信息素养的信息人才的核心知识。信息组织历史悠久，最早来源于图书馆的纸质文献的管理领域；信息构建是一个新内容，它来源于从建筑学的视角来解决信息组织和利用的问题，发展于互联网环境下的信息组织管理领域；知识组织是信息组织发展的高级阶段，是高层次的信息组织和利用。本书将这三个方面的内容有机地整合在一起，兼顾传统信息组织知识及其在网络时代的发展趋势，关注传统的信息组织的原理、方法、流程、技术以及现代的数字资源的微观组织和宏观整合；关注作为信息组织新形式的信息构建的原理、方法、技术和应用；关注作为信息组织高级阶段的知识组织的原理、方法和技术。本书内容丰富，叙述深入浅出，适合作为信息管理与信息系统、图书馆学与情报学、信息资源管理等专业的教学用书，也可作为政府、企业、事业单位的培训用书。

前言

信息组织与信息构建

P R E F A C E

本书由四个部分组成，第一部分为绪言，主要论述信息组织的基础和时代背景，由第一章和第二章构成。第一章论述信息和信息组织的内涵及意义，第二章论述网络环境下的信息组织。第二部分为信息组织，主要论述信息组织的原理、方法、流程、技术和数字资源组织，由第三章、第四章、第五章、第六章共四章构成。第三章论述信息组织的基本原理和方法，第四章论述信息组织的流程与系统设计，第五章论述数字资源的微观组织，第六章论述数字资源的宏观整合。第三部分为信息构建，主要论述信息组织与信息构建关系，信息构建的原理、过程和方法，由第七章、第八章、第九章共三章构成。第七章论述信息组织与信息构建，第八章论述万维网的信息构建，第九章论述信息构建的原理、过程和方法。第四部分为知识组织，主要论述知识组织的原理、方法、技术和应用，由第十章、第十一章构成。第十章论述知识组织的原理、方法和技术，第十一章论述知识组织的应用。各个模块之间既相互关联又相对独立，这样做的目的是为了使用本教材的教师可以根据学习的需要选择其中的模块来教学。

本书各章节的具体分工是：第一章、第二章、第七章、第八章、第九章由周晓英撰写，第六章、第十章、第十一章由马文峰撰写，第三章、第四章、第五章由齐虹撰写。周晓英负责大纲的设计和各章完成后内容的前后统合，各章作者负责本章内容的成文、内容的调整和提出参考引文注释。另外，博士研究生侯人华参与了全书稿的格式整理和校对工作。

关于信息组织、信息构建与知识组织之间的关系，本书认为：信息组织、信息构建与知识组织之间存在着密不可分的关

系，信息组织历史悠久，最早来源于图书馆的纸质文献的管理领域；信息构建是一个新内容，它出现不过 30 多年的时间，而万维网的信息构建出现更是不过 10 年时间，它来源于从建筑学的视角来解决信息组织和利用的问题，发展于互联网环境下的信息组织管理领域；知识组织是信息组织发展的高级阶段，是高层次的信息组织和利用的问题；信息构建和知识组织的本质和核心是信息组织。目前，在信息组织、信息构建与知识组织三部分内容中，除了信息组织有着比较成熟的理论体系之外，信息构建和知识组织都还比较缺乏成熟的理论，把三部分内容比较完整地写进教材中，在国内的同类教材中还没有先例。

本书稿在成文过程中参考引用了大量国内权威的著作和论文，特别是信息组织部分更是借鉴了已有的相对权威的教材，在此，向这些著作的作者表示感谢。

本书的疏漏和不当之处在所难免，恳请业界同人和读者们批评指正！

周晓英

2010 年 8 月 30 日

CONTENTS

信息组织与信息构建

目 录

第一部分 绪言：信息组织的基础和时代背景

第一章 信息、信息困境和信息组织	3
第一节 信息的本质和内涵	4
第二节 信息困境和信息组织的必要性	6
思考题	10
推荐读物	11
第二章 网络时代的信息组织	12
第一节 网络世界信息组织的挑战和目标	13
第二节 网络世界中的信息关系	17
第三节 网络世界发现和组织信息的方式	23
思考题	26
推荐读物	26

第二部分 信息组织

第三章 信息组织的基本原理和方法	29
第一节 信息组织的含义	30
第二节 信息组织的发展	36
第三节 信息组织的基本方法	41
思考题	44
推荐读物	45

第四章 信息组织流程与系统设计	46
第一节 信息组织流程与系统功能	47
第二节 主题分析	48
第三节 信息描述与标引	50
第四节 分类法系统	60
第五节 主题法系统	68
思考题	80
推荐读物	80
第五章 数字信息资源的微观组织	82
第一节 数字信息资源组织概述	83
第二节 数字信息资源编码标准	85
第三节 元数据及其描述标准	97
思考题	106
推荐读物	107
第六章 数字资源的宏观整合：信息整合	108
第一节 信息整合的背景	109
第二节 信息整合的含义	111
第三节 信息整合的特征	116
第四节 信息整合的表现形态	117
思考题	148
推荐读物	148

第三部分 信息构建

第七章 信息构建与信息组织	151
第一节 信息构建的含义和产生背景	152
第二节 信息构建理论和实践的国内外发展状况	160
第三节 信息构建产生的深层原因分析	164
第四节 信息构建与信息组织	169
思考题	182
推荐读物	182

第八章 万维网的信息构建	183
第一节 信息构建的对象——信息集合	184
第二节 万维网的信息构建特点	186
第三节 信息构建的核心要素	187
第四节 万维网信息构建的应用目标	193
第五节 万维网信息构建案例分析	195
思考题	202
推荐读物	202
第九章 万维网信息构建的原理、过程、方法和工具	203
第一节 信息构建的原理	204
第二节 万维网信息构建的一般过程	209
第三节 信息构建的基本方法	213
第四节 信息构建的工具和手段	218
思考题	222
推荐读物	222

第四部分 知识组织

第十章 知识组织原理、方法与技术	225
第一节 知识组织原理	226
第二节 知识组织方法	232
第三节 知识组织技术	252
思考题	261
推荐读物	261
第十一章 知识组织的应用：知识整合	263
第一节 知识整合概述	264
第二节 知识整合的应用——微观整合	265
第三节 知识整合的应用——宏观整合	275
思考题	283
推荐读物	284

**第一部分 绪言：
信息组织的基础和时代背景**

CHAPTER ONE

第一章

信息、信息困境和信息组织

【概 要】

本章分为两节。第一节讨论了信息的概念以及数据、信息、知识和智慧之间的关系；第二节讨论了人类历史上的两次信息爆炸，以及信息爆炸带来的人类知识获取的困难，而信息组织可以成为解决人类知识困境的一种方法，同时阐述了信息组织的必要性。

【学习要点】

信息的含义；数据、信息、知识与智慧之间的关系；人类历史上的两次信息爆炸；信息组织的必要性的体现。

【学习目标】

认识信息的含义；掌握数据、信息、知识与智慧之间的关系；了解人类历史上的信息爆炸的表现；理解信息组织的必要性。

第一节 信息的本质和内涵

一、信息的概念

信息是我们这个时代使用频率最高的词汇之一。围绕着信息出现的事物可以说是丰富多彩、五花八门，如果我们用百度搜索引擎搜索“信息”，网页上会显示“百度一下，找到相关网页约 100 000 000 篇，用时 0.002 秒”。百度百科中也给信息下了三个定义：“（1）信息是物质、能量、信息及其属性的标识。[2006 年，医学信息（杂志）]。（2）信息是确定性的增加。（3）信息是事物现象及其属性标识的集合。”^①如果用 Google 搜索引擎搜索“information”一词，网页上会显示“约有 3 190 000 000 项符合 information 的查询结果，搜索用时 0.05 秒”（2009 年 3 月 12 日）。这样的结果，至少反映了两个问题，一是信息很多很多，只要有条件，只需要不到 1 秒，你就可以得到它；二是网络上关于信息的内容很多很多，它是个相当热门的问题。

信息的概念古已有之，只是在近几十年它才真正成为重要的词汇，人们对信息的认识也在逐步深入。由于人们从不同的侧面、不同的层次解释信息的特征与性质，因而也给信息下了很多的定义。据不完全统计，信息的定义有 100 多种，其中不乏经典的定义，其中代表性的有以下几种。

（1）哈特莱（R. V. L. Hartely）的定义：信息是选择的自由度。

（2）申农（C. E. Shannon）的定义：信息是用来消除和减少随机不定性的东西。

（3）维纳（Norbert Wiener）的定义：信息是人们在适应外部世界，并使这种适应反作用于外部世界的过程中，与外部世界进行互相交互的内容的名称。

（4）朗高（G. Longo）的定义：信息是反映事物的形成、关系和差别的东西，它包含于事物的差异之中，而不在事物本身。

（5）信息是一种可供开发和利用的并具有自身价值的资源。

除了上述经典定义之外，也有类似上述百度百科中来自平民的、非正式的定义。

简单地说，信息是客观事物的各种表现和反映，世界上一切物质形态和物质

^① 百度百科，<http://baike.baidu.com/view/1527.htm>。

运动过程都可以产生信息和信息的运动过程。从无机界到有机界、从宏观的宇宙天体到微观的基本粒子、从自然界到人类社会，种种物质形态都可以成为信息源，都能产生和发出信息。物体在运动过程中，必然伴随着信息的产生、获取、传递、交换、存储等过程，所以任何物体的运动过程都离不开信息的运动过程。从本质上说，信息不是物质的、不是能量的、更不是精神的，它普遍存在于自然界、生物界和人类社会，是事物的属性、内在联系和含义的表征，这种表征在不同的物质运动过程中表现为不同的现象形态，呈现为由低级到高级逐步递进的阶梯和系列。

信息的属性也有很多种，主要包括普遍性、客观性、无限性、相对性、抽象性、依附性、动态性、异步性、共享性、存储性、可传递性、可变换性、可转化性、可伪性等。

二、数据、信息、知识、智慧

虽然任何物质都可以产生信息，但是信息却与人有着最密切的关系，人要生存除了满足基本的生活条件之外，还必须有信息。因此，人所认识的信息应该与认知过程相关，必须把信息放到人的认知阶梯中来认识。

人的认知过程存在着如下认知阶梯（见图 1—1），有学者把它称为情报过程链或信息管理链。^①

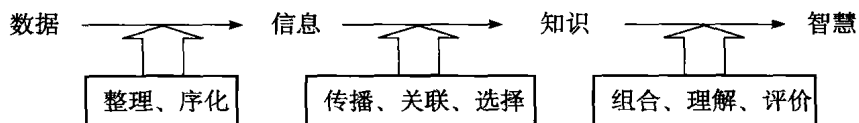


图 1—1 认知阶梯示意图

认知过程实际上是从数据到智慧间的一个连续过程，这个过程各个阶段间的界限不是很清晰，在实际使用中，这几个概念经常被混合或交换使用，原因也许就在于：数据、信息、知识原本就是同一种东西，只是在不同的场合具有不同的特征。比如，本质上无意义的数据在与某种事物伴随出现时就成为信息，信息只要与特定的事物发生关联就可以转化为知识。因此，我们看到，在实际操作中，许多对数据、信息、知识的定义都是用其中一种来定义另一种，或者人们干脆将它们统称为信息或知识。因此，为了说明问题的方便，我们把该认知阶梯中

^① 参见汪冰等：《情报学基础理论研究进展》，见张力治主编：《情报学进展》，第3卷，16页，北京，航空工业出版社，1999。

的数据、信息、知识和智慧统称为认知元素，并对它们做一个限定：

(1) 数据是在本质上没有含义的认知元素。

(2) 信息是有背景的数据，或是通过对数据的排列和表达后，能发现隐藏在数据背后的意义的认知元素。

(3) 知识是发现信息的特定性的关联，是人与信息进行交互后形成的一种认知元素。

(4) 智慧是对知识的理解和运用的内在的认知元素。

也有学者对数据、信息和知识之间的关系做了以下总结：

数据 + 背景 = 信息

数据是信息的原材料

信息 + 经验 = 知识

信息是知识的原材料^①

知识的获取只能通过学习和体验（实践），人的认识能力和理解能力为数据转化为信息、信息转化为知识、知识的有效融合、新知识的创造提供了必不可少的条件。

对于个人而言，人们为了成为有智慧的人，必须要存储数据和信息，并期望能够在恰当的时候把它们转换为知识；人们还需要不断地通过学习，成为能够有效获取和掌控知识的智者。所以对信息的追求是人的本性，是人的高级需求。对于社会而言，信息是社会的记忆，是社会状态的反映，社会的发展与进步要依赖于摄取、管理和利用信息，依赖于信息资源的开发和利用。

第二节 信息困境和信息组织的必要性

一、人类历史上的两次信息爆炸

从 20 世纪 40 年代中期开始，科学技术的迅猛发展带来了科学研究成果的快速的增长，表现形式为各种印刷媒体的文献数量快速膨胀，带来了所谓“信息爆炸”的后果，这是人类历史上第一次信息爆炸。从 20 世纪 90 年代开始，计算机技术和网络技术的发展带来了数字信息的快速增长，除了纸质媒体的信息仍然在

^① 马费成等：《信息管理学基础》，8~9 页，武汉，武汉大学出版社，2002。

增长之外，数字信息的增长速度更是令人惊奇，可以说这是人类历史上第二次信息爆炸。

以下情况可以反映第二次信息爆炸中信息数量的巨大。

从纸质媒体的文献数量看，有数据统计，全世界每年出版的图书和期刊共有约 120 万种，加上其他类型的文献，纸本文献的资料共有超过 500 万种；从出版情况看，世界上平均不到一分钟就有一本新书面世，每年发表论文约 500 万篇。^①

从数字信息数量看，据国际著名的数据交换公司 IDC 统计，仅 2006 年，全球产生、获取和复制的数字信息总量就高达 $1\ 288 \times 1\ 018$ 比特，是有史以来出版的图书信息总量的 300 万倍；每天约有 40 亿个信息单位的信息量发送，预计到 2011 年将达到 1.8 万亿吉比特。

从信息发送的主体看，除了全球范围内的组织机构，主要包括政府、企业、教育机构、研究机构、国际组织、协会学会在通过它们的网站发布着大量的信息之外，由于互联网给个人发布信息提供了有史以来最大的便利和自由，普通的网民也开始越来越多地在网络上发布信息，从严肃的学术研究成果到个人观点、个人情感、信息线索、讨论议论等无所不包，网民提供的信息量越来越大，IDC 预计，2010 年，数字宇宙中有近 70% 的信息由个人建立（IDC，2007）。

对于中国而言，互联网的发展给我们的信息获取和利用带来了巨大的变化，这种变化是我们分析人与信息之间的接触所不能忽视的，下面我们以中国互联网络信息中心（CNNIC）的调查数据为例来说明。^②

（一）中国的网民数量已经跃居全球第一，互联网普及率超过全球平均水平

据 CNNIC 统计，截至 2008 年年底，中国网民规模达到 2.98 亿人，较 2007 年增长 41.9%，互联网普及率达到 22.6%，略高于全球平均水平（21.9%）。^③ 该数据表明，随着我国经济的快速发展、信息化战略的逐步实施、通信和网络技术的发展，互联网的吸引力、网络价值的提升带来了互联网网民数量的稳步增长，中国已经成为互联网大国，接收数字信息成为很多中国人生活的一部分。

（二）中国的互联网基础设施保持快速增长

据 CNNIC 统计，2008 年中国互联网基础资源保持快速增长，IPv4 达到

① 参见 <http://202.204.190.1/learning/jiangyi/ma/%BB%C5%C2%DB.ppt>。

② 参见中国互联网络信息中心：《第 23 次中国互联网络发展状况统计报告》，<http://www.cnnic.cn>。

③ 参见 <http://www.internetworldstats.com>；对比的其他国家和地区互联网普及率为 2008 年 6 月底的数据。

181 273 344个,比2007年增长34%;域名总量达到16 826 198个,比2007年增长41%;域名注册者在中国境内的网站数达到287.3万个,比2007年增长91.4%;网页总数超过160亿个,比2007年增长90%;网络国际出口带宽达到640 286.67Mbps,比2007年增长73.6%。该数据表明,互联网基础设施受到高度重视。由于网民对互联网持续增长的需求,互联网基础设施也在逐步完善和快速发展中。

(三) 互联网成为网民生活中的重要媒体

据CNNIC统计,2008年网民不仅具有巨大的规模,网民总体花费在互联网上的时间也比2007年大大增加,家庭和网吧是网民上网最主要的场所,电脑和手机都可作为上网的设备。该情况表明,继书报、电视之后,互联网正日益成为网民日常生活中的重要媒体。

(四) 网络应用在我国全面展开

据CNNIC统计,2008年网络媒体的使用率较2007年提升了近5个百分点,达到78.5%;用户群体增长7 900万,达到23 400万人;全年搜索引擎用户增长了5 100万人,年增长率达到33.6%;电子邮件使用率为56.8%,与2007年持平;即时通信业务的使用率为75.3%,比起2007年年底,用户群规模增长了5 300万,但使用率降低了6.1%;交友网站较2007年有较大规模的增长,使用率达到19.3%;在我国2.98亿网民中,拥有博客的网民比例达到54.3%,用户规模为1.62亿人;网络游戏用户使用比例从2007年的59.3%升至2008年的62.8%;使用网络音乐的网民比例为83.7%,比2007年的用户数量增长了6 700万人;网络视频用户达到2.02亿,相比2007年年底净增4 000多万用户;网络购物用户人数已经达到7 400万人,年增长率达到60%;网上银行使用率为19.3%;网络炒股应用比例为11.4%;网上教育的使用率为16.5%。该数据表明,各种网络应用在我国已经全面展开。

(五) 互联网对网民的影响

据CNNIC统计,69.3%的网民认为,网上办事减少了他们很多亲临实地的麻烦;64.6%的网民表示,遇到问题时,他们首先会去网上找答案;82.5%的网民认为,互联网加强了他们与朋友的联系;29.0%的网民认为,互联网减少了他们与家人相处的时间;27.6%的网民认为,在网上进行交易是安全的;76.9%的网民认为,上网以后,他们比以前更加关注社会事件。这些数据表明,互联网对于大多数网民而言,是他们的生活助手、获取信息的首选渠道、交往和参与社会活动的重要工具。