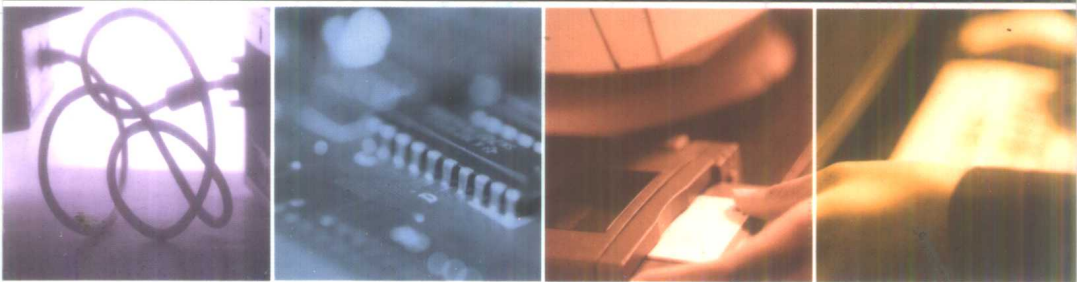


Wangluo Xinxi Ziyuan de Zuzhi



# 网络信息资源的组织

刘嘉 / 著

从信息组织到知识组织

Cong Xinxi Zuzhi Dao Zhishi Zuzhi



北京图书馆出版社

网络信息资源的组织：从信息  
组织到知识组织

刘嘉 著

北京图书馆出版社

**图书在版编目(CIP)数据**

网络信息资源的组织:从信息组织到知识组织/刘嘉著. —北京:  
北京图书馆出版社, 2002. 4

ISBN 7-5013-1894-8

I. 网… II. 刘… III. 计算机网络—信息管理 IV. G203

中国版本图书馆 CIP 数据核字(2002)第 020069 号

**书名** 网络信息资源的组织:从信息组织到知识组织  
**著者** 刘嘉 著

---

**出版** 北京图书馆出版社 (100034 北京西城区文津街7号)

**发行** (010)66126153 传真(010)66174391

**E-mail** Btsfxb@publicf.nlc.gov.cn

**Website** www.nlcpress.com

**经销** 新华书店

**印刷** 北京双桥印刷厂

---

**开本** 850×1168 毫米 1/32

**印张** 11

**版次** 2002年4月第1版 2002年4月第1次印刷

**字数** 263(千字)

**印数** 1—3000

---

**书号** ISBN 7-5013-1894-8/G·507

**定价** 20.00 元

世上活着的事物

都在寻找美好的东西。

卡尔·波普尔

## 前 言

本书是有感于网络环境下的信息资源庞大因而更需要有效组织而写的。网络作为一种新的传播媒体出现以来,网络信息资源数量增加之迅速以日新月异都不足以形容。可是,面对这样海量的信息,却有人发出“淹没于信息却乏于知识”的感喟,此皆因网络信息资源的组织跟不上网络信息资源增长的步伐之故。在这种情况下,笔者以为,有对网络信息资源组织的理念、模式、方法等进行研究之必要,从而探讨此方面较为合理的发展趋势。

人类有别于动物的本质性差异,便在于人类能够有意识地去发现问题,并积极解决问题,从而推动人类社会持续地向前发展。就目前网络资源组织与整理方面所出现的问题,社会各界中凡有能力者,皆积极应对,提出各自的方案,而这自然也是体现自身价值的良机;同时,他们还在此基础上协同工作,以期建立优质的有序的全球网络信息环境。本书即力求全面反映人类在网络信息资源组织方面所进行的各种努力。全书共分八章。第一章为“信息组织概论”,综述人类组织信息和知识的基本原则和方法,并概述人类在各个领域组织信息的基本方略;第二章便是“网络信息资源的组织”,首先对网络信息资源的概念加以界定,随后说明网络信息资源的类型,继而介绍网络信息资源的特殊性,最后,阐述网络信息资源组织的模式并加以评价;第三章即为“网上信息资源检索系统和工具”,在介绍和评价 Archie 和 Gopher 的基础上,重点探讨万维网的检索系统和检索工具;第四章和第五章分别是“网络信息资源的分类和主题体系”和“网络信息资源编目”,其最初的理念皆源于图书馆组织信息资源的思想和方法,在网络环境下,它

们或得到发展,继续发挥着应用的作用,或得到革新,在网络环境下生发出全新的模式;第六章是“元数据概述”,通过论证,提出元数据是组织网络信息资源最重要的基础,并阐释了元数据的基本概念、类型和特征、作用和应用等;紧接着的第七章“都柏林核心元素集”所探讨的都柏林核心,就是目前使用最广泛的元数据类型之一,介绍了都柏林核心的发展历程、元素和限定词、特色和设计原则以及应用等;最后一章是“网络环境下图书馆的作用”,在阐明网络环境对图书馆的影响后,列举图书馆在网络环境下所能够发挥的重要作用,坚信图书馆仍将长足发展。

借此机会,首先,向恩师周文骏教授于笔者攻读博士学位数年期间的淳淳教授致谢意万千,师恩绵绵,泽惠终生;再者,感激黄浦江畔的父母和弟弟以及陆廷鹤老师和师母,他们始终如一的支持永远是笔者力量之源泉;同时,还要向海峡对岸的前辈和师友所给予笔者的莫大帮助致谢意万千。

笔者力求在探讨理论的同时,揭示实践之内涵,从而使本书既有理论上的深度,又不乏实践的依托。然而,总难免诸多不足之处,还请各位专家海涵、指正。

路漫漫兮,其修远矣,吾将上下而求索。谨以此自勉。

二〇〇一年盛夏  
于畅春园



# Wangluo Xinxi Ziyuan de Zuzhi

Cong Xinxi Zuzhi Dao Zhishi Zuzhi

责任编辑：李德升 封面设计：马宁



ISBN 7-5013-1894-8



9 787501 318940 >

ISBN 7-5013-1894-8/C·307

定价：20.00 元

# 目 次

图次	(1)
表次	(1)
前言	(1)
第一章 信息组织概论	(1)
第一节 信息与知识	(1)
第二节 信息组织与知识组织	(7)
第三节 不同机构开展的信息组织工作	(13)
第四节 小结	(22)
第二章 网络信息资源的组织	(24)
第一节 网络信息资源的界定	(24)
第二节 网络信息资源的类型	(26)
第三节 网络信息资源的特殊性	(31)
第四节 网络信息资源的组织模式及其评价	(41)
第五节 小结	(51)
第三章 网上信息资源检索系统和工具	(55)
第一节 Archie系统和Gopher系统	(55)
第二节 万维网	(63)
第三节 小结	(74)
第四章 网络信息资源的分类和主题体系	(80)
第一节 网络信息资源的分类体系	(81)
第二节 网络信息资源的主题体系	(100)
第三节 小结	(112)



第五章	网络信息资源编目	(117)
第一节	图书馆以 MARC 格式为网络信息资源 编目的原因	(118)
第二节	网络信息资源编目项目	(125)
第三节	用于网络信息资源编目的美国机读目录格式	(135)
第四节	小结	(166)
第六章	元数据概述	(170)
第一节	何为元数据	(170)
第二节	元数据是全新的概念吗	(176)
第三节	元数据的类型与特征	(180)
第四节	元数据的作用	(186)
第五节	元数据的应用	(189)
第六节	利用元数据进行知识管理	(194)
第七节	与元数据相关的其它问题	(196)
第八节	小结	(201)
第七章	都柏林核心元素集	(206)
第一节	都柏林核心的发展历程	(207)
第二节	都柏林核心元素集的元素和限定词	(220)
第三节	都柏林核心的特色和设计原则	(232)
第四节	都柏林核心的应用	(236)
第五节	都柏林核心与机读目录格式之间的关联	(239)
第六节	都柏林核心与机读目录数据的映射	(242)
第七节	小结	(254)
第八章	网络环境下图书馆的作用	(259)
第一节	网络环境对图书馆的影响	(260)
第二节	网络环境下图书馆的作用	(269)
第三节	网络信息资源组织与图书馆信息资源组织 之间的关系	(278)

第四节 小结 .....	(282)
附录 1 使用分类体系组织网络资源的网站 .....	(287)
附录 2 三种元数据格式的数据元素和数据结构 之间的映射 .....	(295)
附录 3 网络信息资源编目实例 .....	(302)
附录 4 USMARC 记录、DC HTML 和 DC RDF 记录对照实例 .....	(310)
附录 5 书中常用术语中英对照表 .....	(317)
附录 6 中国大陆与台湾网络资源组织方面用语对照表 .....	(320)
参考与引用文献 .....	(324)

## 图 次

图 1-1	影响知识管理活动的知识的属性 .....	(4)
图 1-2	知识的纵向增值 .....	(6)
图 1-3	知识的横向增值 .....	(7)
图 1-4	一网上数据库的 MARC 记录实例 .....	(15)
图 1-5	中国大陆档案著录文件级格式 .....	(16)
图 1-6	中国大陆档案著录案卷级格式 .....	(17)
图 1-7	美国 The Sampler Museum 的卡片目录格式 .....	(20)
图 2-1	PURL 的主要组成部分 .....	(38)
图 2-2	PURL 与 URL 之间的关系 .....	(39)
图 2-3	高级系统图 .....	(39)
图 2-4	信息系统中数字对象的生命周期 .....	(49)
图 3-1	Yahoo! 的主页 .....	(67)
图 3-2	AltaVista 的检索界面 .....	(68)
图 3-3	Dogpile 的主页 .....	(70)
图 4-1	WebDewey 记录的结构 .....	(86)
图 4-2	NetFirst 的高级检索界面 .....	(92)
图 4-3	“雅虎中国”的大类体系图 .....	(96)
图 4-4	“雅虎中国”中“参考资料”的下位类 .....	(96)
图 4-5	Infoseek Education Topics .....	(98)
图 4-6	CyberHound Expert Search 浏览结果界面 .....	(102)
图 4-7	OMNI 使用 NLM 的主题列表 .....	(105)
图 4-8	OMINI 字顺主题列表:生态学 .....	(106)
图 4-9	Google 的基本搜索界面 .....	(110)

图 4-10 Google 的高级搜索界面 ..... (111)

图 8-1 美国 1994、1996 年已经与因特网联网和 1997 年  
计划要与因特网联网的公共图书馆占全美  
公共图书馆之比例 ..... (265)

## 表 次

表 1-1	信息和知识的区别 .....	(6)
表 2-1	Archie、Gopher 组织和整理网络信息资源的方式对照表 .....	(43)
表 4-1	数据库规模统计数字和主题组织属性 .....	(88)
表 4-2	利用 DDC 编码将因特网资源划分为十大类 .....	(89)
表 4-3	NetFirst 详细记录样例 .....	(90)
表 4-4	国内外几种主要搜索引擎的大类体系 .....	(93)
表 4-5	Yahoo Education Topics .....	(97)
表 4-6	“网络指南针”的分类目录 .....	(99)
表 4-7	取自 EELS 的记录样本 .....	(108)
表 5-1	应用的 AACR2R 规则与 USMARC 格式 字段对照表 .....	(138)
表 5-2	856 字段所包含的单元 .....	(164)
表 6-1	不同类型的元数据及其功能 .....	(181)
表 6-2	元数据的属性和特征 .....	(182)
表 6-3	元数据格式的类型 .....	(184)
表 6-4	各组元数据格式的特征 .....	(185)
表 6-5	知识的增值过程 .....	(196)
表 7-1	都柏林核心历届会议简表 .....	(207)
表 7-2	对都柏林核心 15 个元素的归类 .....	(221)
表 7-3	台湾都柏林核心元素集元素译名 .....	(222)
表 7-4	DCMI 公布的限定词 .....	(231)
表 7-5	都柏林核心和 UNIMARC 数据单元对照表 .....	(245)

表 7-6 都柏林核心到 USMARC 的对照表…………… (249)

表 8-1 与因特网的连接为公共图书馆带来的显著益处  
…………… (266)

# 第一章 信息组织概论

组织是人类最基本的本能之一。根据精神病学家的观察,婴儿将图像组织、区分为诸如“面孔”或“事物”最简单的类别,儿童在游戏的过程中也会有一些组织方面的行为;而在成人世界里,组织行为更是缺之不可,否则这个世界就会变得一片混沌。对于一些组织欲望特别强烈的人而言,他甚至只有在工作环境完全进入一种有序状态之后,才可能开展工作。毋庸置疑,人类的认知(Learning)就是建立在分析和组织数据(Data)、信息(Information)和知识(Knowledge)的能力之基础上的。

## 第一节 信息与知识

信息在人类社会中无处不在,它与物质、能量被视作为维系人类社会生存与发展的必不可少的三要素,也是人类社会的宝贵资源和财富;而知识乃人类文明的产物,是人类社会实践经验的总结,是人对主观世界的反映与认识的结晶。

### 1 何为信息与知识

信息一词的拉丁词源是 Information,意思是通知、报道或消息。在中国历史资料中,信息最早出自唐诗,是音信、消息的意思。一直沿用到二十世纪中叶,其科学含义才被逐渐揭示出来。<sup>1</sup>而对信息深入的科学研究,还是始于西方社会。1928年,R. V. Hartley发表《信息传播》(Transmission of Information),区分了消息和

信息,指出“信息是指有新内容、新知识的消息”。1948年,N. Wiener发表《控制论》(Cybernetics),从更广阔的领域研究信息,在更开阔的意义上定义了信息,认为信息是“我们在适应外部世界、控制外部世界的过程中同外部世界交换的内容的名称”。同年,C. E. Shannon出版《通信的数学理论》(Mathematical Theory of Communication),提出了信息量的概念和信息熵的计算方法。<sup>2</sup>

显然,关于信息的概念,在图书情报领域、计算机与通讯科学领域、经济管理领域等各个领域都有不同角度的解释。颇具权威性的《中国大百科全书》中有关“信息”的条目提出:“按照狭义的理解,信息是用来消除不定性的东西。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;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。正因为信息是事物运动的状态和方式(本体论意义),是关于事物运动状态和运动方式的反映(认识论意义),它才可以用来消除人们认识上相应的不定性。这是广义理解和狭义理解之间的内在联系。”<sup>3</sup>

根据英国科学家 K. Popper 在有关“三个世界”的理论,信息可分为三大类。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,即主体论或认识论意义上的隐性信息,它反映人类所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的概念世界的信息,即主体论或认识论意义上的显性信息,它反映人类所表述的事物运动状态及其变化方式,用语言、文字、图像、影视、数据等各种载体来表示,汇合成一个实在的自主的“信息世界”。“三个世界”理论通过类分方法将信息划分为独立的三大部分,得以使人们全面、准确地认识和把握信息。<sup>4</sup>

信息之于人类的最重要意义,就在于它在人类认识和实践中的



所发挥的特殊作用。人类认识世界的过程,实际上就是不断从外界取得信息和加工信息的过程;而人类改造世界的过程,则是把加工外部信息所得到的“主观”信息(表现为方针、政策、计划等等)反作用于外部世界的过程。可见,没有信息,就不能认识世界;没有信息,也不能有效地改造世界。

而知识是比信息更高层次的人类的精神财富。知识是人们对自然界、人类社会和思维活动规律的认识和掌握,是人们把大量事实和思想投入熔炉后经提炼和组合而成为可供使用的东西。知识一般可用“6W”和“1Q”来概括,即“知道什么”(know - what)、“知道为什么”(know - why)、“知道怎么做”(know - how)、“知道谁”(know - who)、“知道何时”(know - when)“知道何地”(know - where)和“知道数量”(know - quantity)。<sup>5</sup> 知识包括事实、原理、概念和能力。除了一部分基础知识外,大部分是专门知识,是关于某个领域、某个主题、某种思想方法、某门科学、某种技术、某种价值体系、某种社会组织形式的认识。知识本身或者就是科学技术,或者可以转化为科学技术,而科学技术则是推动社会发展的第一生产力。

知识是人们通过实践对客观事物及其运动过程和规律的认识,是信息的一部分,即被人们理解和认识并经头脑重新组织和系列化的那部分信息,是经验、技能的总结,是组织发展竞争力的基础,在反复实践和认识的过程中,人脑通过对相关概念的判断、组合和推理,形成对事物本质的认识,构成头脑中的知识,可称为主观知识,即隐性知识(Tacit knowledge);如果经过各种载体表达出来,则成为客观知识,即编码型的知识(Encoded knowledge)。依照对于信息的分类方法,这两类知识分别对应于“三个世界”理论的第二类和第三类信息。而第一类信息不能直接形成知识,这也说明了知识是信息的一部分。<sup>6</sup>

影响知识管理活动的知识的属性包括知识被掌握的形式、知