

ZHUTITU
ZILUN YU YINGYONG
FANGFA YANJIU

主题图理论与 应用方法研究

——以四川省阿坝州旅游文化信息资源组织为例



◆ 李清茂 著



四川大学出版社
SICHUAN UNIVERSITY PRESS

ZHUTITU
LILUN YU YINGYONG
FANGFA YANJIU

主题图理论与 应用方法研究

——以四川省阿坝州旅游文化信息资源组织为例

◆ 李清茂 著



四川大学出版社
SICHUAN UNIVERSITY PRESS

责任编辑:曾 鑫

责任校对:李金兰

封面设计:李金兰

责任印制:李 平

图书在版编目(CIP)数据

主题图理论与应用方法研究:以四川省阿坝州旅游文化信息资源组织为例 / 李清茂著. —成都:四川大学出版社, 2011. 7

ISBN 978-7-5614-5390-2

I. ①主… II. ①李… III. ①信息管理—研究 IV. ①G203

中国版本图书馆 CIP 数据核字 (2011) 第 151637 号

书名 主题图理论与应用方法研究
——以四川省阿坝州旅游文化信息资源组织为例

著 者 李清茂
出 版 四川大学出版社
地 址 成都市一环路南一段 24 号 (610065)
发 行 四川大学出版社
书 号 ISBN 978-7-5614-5390-2
印 刷 郫县犀浦印刷厂
成品尺寸 185 mm×260 mm
印 张 14
字 数 313 千字
版 次 2011 年 7 月第 1 版
印 次 2011 年 7 月第 1 次印刷
定 价 30.00 元

◆读者邮购本书,请与本社发行科
联系。电 话:85408408/85401670/
85408023 邮政编码:610065

◆本社图书如有印装质量问题,请
寄回出版社调换。

◆网址:www.scupress.com.cn

版权所有◆侵权必究

- 四川省科技厅 2010 年科研立项支持课题：“阿坝州旅游客户行为分析与推荐系统实现”，课题编号：2010JY0J41
- 四川省教育厅科研基金 2010 年重点立项支持课题：“主题图理论与实践探索”，课题编号：10SA090
- 阿坝师范高等专科学校 2011 年校级重点立项支持课题：“主题图与传统信息组织方法的比较研究”，课题编号：ASA11-30
- 阿坝师范高等专科学校“网络与智能信息处理”校级科研团队支持项目

序

党跃武

(四川大学公共管理学院教授, 教育部图书馆学学科教学指导委员会委员)

当今社会是一个充满幻想的时代, 同时又是一个追求效用的时代。美国哈佛大学信息政策研究中心主任安东尼·G·欧廷格曾经指出: “没有物质, 就什么东西也不存在; 没有能量, 就什么事情也不发生; 没有信息, 就什么东西也无意义。” 换句话说, 缺少物质的世界是一个空虚的世界, 缺少能量的世界是一个死寂的世界, 缺少信息的世界将是一个混乱的世界。

信息资源: 数量的爆炸和需求的迸发

当年丹尼尔·贝尔提出所谓“信息爆炸”和约翰·奈斯比特感叹“我们现在大量生产信息, 就像我们过去大量生产汽车一样”的时候, 信息、信息资源的生产和利用还远不能够与网络时代同日而语。在世纪之交, 据专家估计, 20 世纪 40 年代以来产生和累积的信息量已经大大超过了在这之前人类有史以来的所有信息量之和。19 世纪以来, 人类知识信息每 50 年增长一倍, 20 世纪中叶每 10 年增长一倍, 20 世纪 70 年代以后每 5 年增长一倍。信息来源于社会的各个角落, 信息流向社会的各个层面。我们如同海洋中的动物在信息之海中游弋。

信息无处不在, 俯仰皆拾。人们的信息需求同样在迸发, 人们对信息的渴望从来没有像今天这样强烈, 人们的学习、生活和工作都离不开信息的参与。信息需求的自觉和自发地显现, 充分说明社会信息意识——对信息功能的认同意识, 对信息来源的选择和开发意识, 对信息内涵的同构和再生意识, 都得到了充分的加强。因此, “我们都被淹没在信息的海洋中, 但是我们渴求新的信息”。人们的信息需求一方面表现出: 我们生活的环境日趋复杂化, 需要我们更加充分地认识和利用相关因素, 使我们的行为更加符合社会发展规律; 另一方面表现出: 为了处理和分析日趋复杂的工作对象, 我们只有如饥似渴地追求信息、获取信息, 才能进一步促进信息的利用和再利用。同时, 人们的信息需求动机不再简单地体现为学习、研究和娱乐, 现代社会的发展使信息用户的信息需求的目的趋向于多元化, 求知、启发、共鸣、参考、决策、获利、猎奇、觅新、享乐、休闲等多元化需求使信息需求及其满足更加复杂化。

信息迷宫: 利用的困惑和服务的困境

信息资源的爆炸和信息需求的迸发结伴而行。于是, 我们深深地陷于一个矛盾的环境之中。我们冷静地看到, 在信息管理和利用中, 一方面信息数量在绝对激增, 另一方

面信息质量呈现相对贫乏状况。人类社会中漫天飞舞的信息并非对于每个人都具有等同价值，而最让人困惑的是——有用信息和无用信息的混杂导致信息质量的相对贫困。信息数量总是在不断累积，然而在信息质量方面存在的问题却并不那么乐观。第一，信息数量增长的同时，信息却在不断老化，信息的老化使得信息总量中有效信息的比率并不随着信息数量增长而绝对地同步上升。第二，信息数量增长的过程中，有用信息和无用信息都在增长之中，也就是说，增长的信息中有相当一部分是无用信息。第三，即使是有用信息中，不同质量级别的信息也有不同的增长速度，质量级别越高的信息数量越少，增长速度越慢。的确，“我们都被淹没在信息的海洋中”。

我们冷静地看到，在信息管理和利用中，一方面存在着信息不虞现象，另一方面却出现信息相对过剩。曾经有一位叫霍雷斯·瓦尔坡的作家描述了一个古怪的岛国“不虞国”。在那里，你想要的东西老是找不到，你不要的东西偏偏会碰上。在人们利用信息的过程中由于缺乏对信息的合理和科学的控制手段和方法，如同“不虞国”人一样，获得的信息不能满足人们的信息需求，人们需要的信息不能有效获取，出现了这样一种矛盾的信息利用现象，即所谓的信息不虞现象。同时，由于信息利用意识的缺乏、信息传递手段的不足、相似信息的互相干扰等原因，人们的信息需求没有得到满足的同时，存在着大量的信息未能得到全面有效的利用而形成的信息提供过剩的虚假状况。

凡此等等，我们终于体会到——信息正在越过国界，越过时间和空间界线，既给予我们一个广阔的天地，又让我们承受一种信息压力，“信息如同麦克斯韦妖”。我们进入了一个我们自己创造的信息迷宫中。

信息技术：引领我们走出信息迷宫

在信息迷宫中，我们有着太多的焦虑。那么，我们如何才能走出信息迷宫？也许我们曾经使用过许许多多的词汇来描绘信息和信息资源的重要性，但是，所有关于信息重要性的描述都不是严格意义上的，都有一定的约束条件——信息不是杂乱无章的，只有经过管理和控制的信息才具有重要性。美国著名社会学家约翰·奈斯比特指出：“信息革命来到后，我们第一次才有一种建立在不仅可以再生而且可以自生的重要资源之上的经济。资源缺乏不成问题了，但被资源淹没倒是问题。”诺贝尔奖得主阿罗和西蒙当年在评说信息高速公路时都不约而同地提出了信息“过滤”问题。阿罗指出：“在‘信息高速公路’体系中，人变成了接待室，成了电脑化的接待室。我们建立它的目的是要使得人类更为有效率。”西蒙也指出：“它不应仅仅是让大量信息涌向你，而应是帮助你去‘过滤’你需要的信息。这才是它的可取之处。”

因此，信息管理和信息交流活动中，我们需要一种信息“过滤”装置，信息重组与分配的装置。这就是科学的信息组织。对于相当多用户来说，困扰他们的不是信息太少，而是信息太多、相关信息太少的问题。信息来源渠道更加多样，物质载体更加丰富，信息内涵更加精深，利用需求更加复杂，以满足信息用户不断增长的信息需求为己任的信息管理工作，必须充分发挥信息组织的功能和作用。信息不断地产生于社会的各个角落，产生于不同的人，产生于不同的学科和领域，产生于不同的信息形态。如果没有经过有效的控制，它就处于一种自流状态之中，信息与信息之间缺乏内在的联系，特

别是各种内容和形式具有极大相似性的信息无法集中、对比，信息与利用它的人的联系也是随机的。“没有经过控制和处理的信息，不仅不是我们的朋友，甚至是我们的敌人。”然而，从人们信息需求的表现形式来看，人们总是易于理解和运用高度组织和系列化的信息。人们的信息需求总是集中于一定时点，即具有一定的学科或领域的范围、一定的时间和空间的约束、一定的形态和方式的选择，因而人们获取信息的行为一般来讲是理性的而不是漫无边际的、具体的而不是杂乱无章的。因此，我们必须对来源不同、形态不一、水平不齐、内容不等的各种信息进行科学的信息组织活动，从而精化信息、序化信息、共享信息。

当然，我们这里特别强调在信息时代和网络时代，信息组织的任务更加艰巨。这并不是说在漫长的信息管理和利用活动中信息组织是什么新生事物。作为信息管理的主要分支，信息组织一直发挥着极其重要的作用。无可置疑的是，现代信息技术的导入，把信息管理推上一个新的发展水平，也把信息组织引入一个新境界，使信息组织以前连想也未曾想到的新概念和新设想变成了现实。信息组织从最完整的意义上实现了整序信息、科学分流、促进选择、保证利用的组织职能。这种组织职能已经绝对不满足于告诉人们有什么信息和怎样找到信息，它不仅可以告诉人们我们这里有什么信息，还可以告诉在其他地方有什么信息；不仅你可以找到你最需要的信息，而且可以知道什么信息适合于你；不仅你可以利用一种途径获得信息，而且可以多途径、多角度查检信息。在这里，信息的概念被高度扩充，它不仅是文献信息，而且是非文献信息；不仅是文字信息，而且是声音、图形或动感信息；不仅是知识信息，而且是满足你非学习研究型需求的其他信息；不仅是提供线索的二次信息，而且是直接满足要求的一次信息和三次信息；不仅是被动接收的信息，而且是你参与生成和选择的信息；不仅是本地直接存取信息，而且是远程存取网络信息……

所有这一切，反映了基于网络环境和网络需求的现代信息组织的前沿和现实性、精细化和普及化、协作性和专业性、标准化和多样化、体系化和特色化、市场性和调控性、自动化和智能化的有机结合。现代信息管理中信息组织活动的联合化、信息组织产品的商品化、信息组织方式的专营化、信息组织成果的媒体化、信息组织基础的需求化成为这一时期最明显的标志。在未来发展的道路上，信息组织的功能不仅不会弱化，而且会进一步强化。简单地说，所谓信息组织即信息序化或信息整序，也就是利用一定的科学规则和方法，通过对信息外在特征和内容特征的表征和序化，实现无序信息流向有序信息流的转换，从而保证用户对信息的有效获取和利用、信息的有效流通和组合。信息组织是信息管理的重要环节和基本工作，是信息资源开发利用的主要手段，是信息传播的前期准备。因此，信息组织对信息的序化和整理可以在一定程度上解决信息普泛性与信息效用个体性、信息利用规律性的矛盾，从而达到甄别信息、精化信息、重组信息的作用。

信息组织：导航仪或者指南针……

在信息管理和利用活动中，信息组织的对象是什么？信息组织扮演着什么样的角色？信息组织必须遵循什么样的法则？信息组织的具体内容是什么？信息组织的运作将

如何进行?信息组织如何利用新型信息技术?网络环境下的信息组织具有哪些新特点?一个又一个问题摆在我们面前,需要我们积极地进行思考,给予准确的、全面的回答。毫无疑问,本书就是我们初步思考和探索的结果。

在这些问题的思考和探索中,我们特别要强调信息技术在信息组织的导入和网络环境与信息组织发展问题。当年,托勒密一世曾经想把世界上所有的图书汇于一个小小的亚历山大图书馆,由它来对所有的图书进行管理和组织,然而古罗马军队的一把火烧毁了“亚历山大神话”。虽然德国人格斯纳又梦想着依靠自身的力量独立编撰一部所谓的《世界书目》,但是第一版《世界书目》收录的文献还不如几乎同时期的《永乐大典》。其实,在当时的技术环境下,“亚历山大神话”和“格斯纳梦想”永远不可能成为现实。今天信息技术的突飞猛进完全有可能在现代信息技术条件下使“亚历山大神话”和“格斯纳梦想”成为“现实”。因此,新技术在信息组织中的导入成为一个重要的热点问题。

在现代网络环境下,信息资源组织的新特点正在逐步地得以体现。多载体的信息成果成为主流,虚拟信息成果更加引人注目,各种类型信息成果相互交融形成的组合式信息成果成为新的集成化组织形式。现行的网络信息检索和信息分享机制的建立,虽然还不能说是一种全新的信息资源组织和利用模式,甚至有的在某种意义上还是萌生于图书馆现行运作机制的原型信息资源组织和利用模式。但是,这些新的技术和手段的确为现代图书情报机构处理和整序新型信息资源提供了很好的条件。网络教室、网络资源指南、网络资源目录、网络搜索引擎、网络资源导航站、WWW数据库、多媒体资源系统、光盘数据库、虚拟资源目录等各种网络信息资源组织成果,正在不断成熟和完善。信息组织成果建立于面向数据的数据驱动技术机制或面向需求的需求驱动技术机制之上,如何在两种技术机制间寻求平衡点,具有越来越强烈的现实需要,也具有越来越明确的技术可行性。伴随着人工智能技术的发展,在知识经济时代,尤其是知识管理异军突起的条件下,建立于知识采掘、知识表示和知识应用等知识管理技术,面向对象数据库技术、超文本和超媒体技术、数字技术等计算机管理技术之上的现代的信息组织中非线性、联想型数据—需求驱动技术机制已经日臻完善和成熟。

我们完全可以说,现代信息技术使信息组织活动如虎添翼。然而,新技术的引进并不绝对排斥信息组织甚至是文献组织的原有技术的进一步发展和利用。在某种意义上,中国传统目录学的“辨章学术、考镜源流”思想就是现代信息组织中主题揭示和内容分析的滥觞。对于我们来说,面对新的环境,包括技术环境、需求环境、资源环境等,一方面要根据新环境提出新课题,寻求新技术解决新问题和以前无法解决的问题;另一方面要对原有技术进行挖潜更新,使成熟技术在新环境下获得“第二个春天”。因此,我们在把握现代信息组织理论、技术和方法的时候必须坚持这一正确的态度。我们要及时引进、消化和吸收现代信息技术并努力解决各种信息资源(包括传统信息资源和现代信息资源)的组织和管理问题。同时,我们的目光也不能够仅仅停留在现代信息技术之上,还必须把传统信息组织技术加以发扬光大。要发挥信息组织在整个信息管理活动中的作用,必须以信息用户需求和信息资源为导向,为促进信息效能的实现,新的信息组织的形式多样化、手段自动化、功能立体化、成果体系化、运行集约化、流程高效化、管理集成化的态势将是不可避免的。

关于主题图和本书

在信息组织技术的发展中,主题图技术已经成为研究和实践的一个热点。我们可以利用一定的知识组织体系,对信息系统资源集合的主题内容结构、主题词汇、主题间相互关系以及主题与具体资源的链接进行描述,形成资源集合的主题图。主题图直接用于信息导航系统的知识化浏览,建立资源集合的主题索引或交叉参照,链接复杂主题范围的分布式资源来建立虚拟知识体系,通过主题概念与资源的不同链接,在同一资源体系上建立面向不同主题体系或不同用户的资源界面。主题图最早可以追溯到20世纪90年代初,其目的是希望建立起智能化电子索引并支持相互间的融合。主题图的前身是“主题导航地图”(Topic Navigation Maps)。为了便于专业信息的组织和顺应跨学科领域的发展趋势,主题图(Topic Maps)的概念得以提出,并于1999年获得国际标准组织的认证,2000年成为国际标准(ISO/IEC 13250:2000 Topic Maps)。2001年3月2日发布XTM(XML Topic Maps)第一版,2001年10月,XTM DTD被ISO13250所采纳。2002年由TopicMaps.Org负责修订第二版(ISO/IEC 13250:2002)。随着主题图技术的不断发展,目前已经出现了一系列专门建立和管理主题图的工具。网络主题图更是广泛应用于网络信息检索、网络信息组织、网络信息挖掘、网络信息资源整合、网络信息过滤和数字资源知识管理等各个方面。

作为李清茂同志的硕士论文指导教师,我有幸成为《主题图理论与应用方法研究——以四川省阿坝州旅游文化信息资源组织为例》一书的第一个读者。我非常高兴地看到,本书主要具有三个特点:第一,研究的完整性。在研究的设计和实施中,本书全面把握相关理论、技术和方法,从基础概念与方法、理论模型与语法、应用工具与开发等方面,较为系统地探讨了主题图相关问题,为进一步深入了解主题图技术提供了翔实的学习读本。第二,研究的实践性。本书突出技术的应用性和可操作性,结合四川省阿坝州旅游文化信息资源组织进行了较为科学的实证分析和实验研究,为进一步开发四川省阿坝州旅游文化资源提供了重要的开发工具。第三,研究的扩展性。本书在旅游文献知识管理模型和信息系统模型构建与实现方面进行了初步的探索,已经形成了较为突出的研究成果,为进一步开展主题图技术的开发应用提供了可以借鉴的宝贵经验。

在信息组织的任务更具挑战性,信息服务的需求更具复杂性,信息管理的手段更具交互性的今天,我们衷心地期待着,并且完全相信,由于李清茂同志这样的青年研究者的加入,在信息组织领域,一定会涌现出理论与实践紧密结合,充分反映学科发展趋势,积极引领技术开发应用的更多、更好的优秀研究成果。

目 录

主题图基础理论篇

第 1 章 绪论	(3)
1.1 主题图发展历程	(3)
1.1.1 主题图引出	(3)
1.1.2 信息知识结构揭示	(4)
1.1.3 主题图发展历程	(8)
1.2 主题图理论研究现状	(11)
1.2.1 国内外研究现状	(11)
1.2.2 国外相关研究资源	(14)
1.2.3 主题图应用案例	(15)
1.3 全书基本框架	(17)
参考文献	(18)
第 2 章 主题图理论基本原理	(19)
2.1 主题图理论内涵与概念辨析	(19)
2.1.1 主题图概念	(19)
2.1.2 主题 (Topic)	(20)
2.1.3 关联 (Association)	(23)
2.1.4 资源指引 (Occurrence)	(26)
2.1.5 主题图的层次关系	(28)
2.1.6 主题图概念模型	(29)
2.2 主题图辅助要素——IFS	(30)
2.2.1 议题身份识别 (Subject Identity)	(30)
2.2.2 主题分面 (Facet)	(31)
2.2.3 主题范围 (Scope)	(32)
参考文献	(34)
第 3 章 主题图与传统信息组织方法比较研究	(36)
3.1 传统信息组织方法分析	(37)
3.1.1 元数据	(37)

3.1.2	主题分类法	(40)
3.1.3	叙词表	(43)
3.1.4	分面分类	(45)
3.1.5	本体	(46)
3.1.6	其他相关方法	(47)
3.2	主题图理论再认识	(47)
3.2.1	议题名称	(48)
3.2.2	主题类型	(49)
3.2.3	主题资源指引	(49)
3.2.4	主题关联	(50)
3.2.5	信息组织方法综合比较	(50)
	参考文献	(52)
第4章	基于主题图实现传统信息组织方法	(55)
4.1	基于主题图实现叙词表与分类法	(55)
4.1.1	利用主题图实现叙词表	(55)
4.1.2	利用主题图实现分面分类法	(57)
4.1.3	利用主题图融合元数据与分类法	(58)
4.1.4	利用主题图融合收益与成本	(59)
4.2	主题图实现技术的其他问题	(60)
4.2.1	主题图对查询的支持	(60)
4.2.2	主题图模式约束	(61)
4.2.3	主题图合并与身份识别	(61)
	参考文献	(62)

主题图模型与语法篇

第5章	主题图数据模型	(67)
5.1	主题图元模型	(67)
5.1.1	什么是元模型	(67)
5.1.2	主题图元模型约束	(68)
5.2	主题图数据模型定义	(70)
5.2.1	主题图数据模型概要	(70)
5.2.2	主题图项目 (Topic Map Item)	(71)
5.2.3	主题项目 (Topic Items)	(72)
5.2.4	主题名称项 (Topic Name Items)	(74)
5.2.5	变量项目 (Variant Items)	(75)
5.2.6	资源指引项目 (Occurrence Items)	(75)
5.2.7	关联项目 (Association Items)	(76)

5.2.8 关联角色项目 (Association Role Items)	(77)
5.3 模型操作与议题标识符	(78)
5.3.1 主题合并操作	(78)
5.3.2 议题标识符定义	(80)
参考文献	(82)
第 6 章 XTM 描述语法规范	(84)
6.1 XTM 语法概述	(84)
6.1.1 XTM 整体架构	(85)
6.1.2 <topicMap>元素	(86)
6.1.3 XTM 文档示例	(86)
6.2 <topic>元素相关语法	(87)
6.2.1 概念辨析	(87)
6.2.2 <topic>元素	(87)
6.2.3 <subjectIdentity>元素及其结构	(88)
6.2.4 <instanceOf>元素及其结构	(90)
6.2.5 <baseName>元素及其结构	(91)
6.2.6 <occurrence>元素及其结构	(94)
6.3 <association>元素相关语法	(95)
6.3.1 <member>元素	(96)
6.3.2 <roleSpec>元素	(96)
6.4 <mergeMap>元素相关语法	(97)
6.4.1 <mergeMap>元素语法规范	(97)
6.4.2 主题合并规则	(98)
6.4.3 主题合并案例	(100)
参考文献	(104)
第 7 章 主题图议题识别机制	(105)
7.1 议题识别概述	(105)
7.1.1 议题 (Subjects) 与主题 (Topics) 概念辨析	(106)
7.1.2 议题识别的意义	(106)
7.1.3 议题可寻址性分析	(106)
7.1.4 议题指示器 (Subject Indicators) 与议题标识符 (Subject Identifiers)	(108)
7.1.5 议题识别实例	(108)
7.2 公开议题识别机制 (Published Subjects)	(110)
7.2.1 议题指示器的缺点	(110)
7.2.2 公开发行人机制	(111)
7.2.3 公开议题工作机制案例	(111)
7.2.4 如何采纳 PSI	(112)

7.3 公开议题识别机制 (PSI) 的要求与建议	(112)
7.3.1 发行 PSI 的要求	(112)
7.3.2 发行 PSI 的建议	(113)
参考文献	(113)

主题图开发工具篇

第8章 主题图应用支持工具	(117)
8.1 TM4J 主题图开发工具	(117)
8.1.1 TM4J 的安装	(118)
8.1.2 TM4J 基本包	(118)
8.1.3 TM4J 基本结构	(120)
8.1.4 TM4J 核心接口	(122)
8.1.5 创建 TopicMapProvider	(124)
8.1.6 TopicMapProvider 配置属性	(126)
8.1.7 TopicMapProvider 相关操作	(128)
8.2 OKS 知识开发组件	(129)
8.2.1 Ontopia 主题图引擎	(130)
8.2.2 Ontopia 导航框架	(131)
8.2.3 Ontopia 主题图展示工具	(131)
8.2.4 Ontopia Web 编辑框架	(132)
8.2.5 Ontopia 主题图编辑工具	(132)
8.2.6 Ontopia 可视化组件	(132)
8.2.7 Ontopia 其他功能	(133)
参考文献	(133)
第9章 Ontopia 主题图引擎工作原理	(135)
9.1 Ontopia 主题图引擎工作机制概要	(135)
9.1.1 OKS 知识组件概要	(135)
9.1.2 Ontopia 引擎结构及运行机制	(136)
9.1.3 Ontopia 引擎核心接口结构	(138)
9.1.4 Ontopia 引擎核心其他接口	(139)
9.1.5 Ontopia 引擎 API 应用支持	(140)
9.2 Ontopia 引擎支持关系数据后台连接	(146)
9.2.1 安装 RDBMS 后台连接	(146)
9.2.2 RDBMS 后台连接器 API	(149)
参考文献	(152)
第10章 主题图查询语言	(154)
10.1 主题图查询语言概述	(154)

10.1.1	与其他标准之间的关系	(155)
10.1.2	结构与语法	(155)
10.1.3	查询语言特征	(155)
10.2	TMRQL 查询语言	(156)
10.2.1	TMRQL 主要关系视图	(156)
10.2.2	TMRQL 功能函数	(158)
10.2.3	TMRQL 特点分析	(161)
10.3	Tolog 查询语言	(162)
10.3.1	Tolog 语言概况	(162)
10.3.2	Tolog 语言高级特征	(165)
10.3.3	运行 Tolog 查询	(170)
	参考文献	(171)

主题图应用方法篇

第 11 章	主题图应用方法研究	(175)
11.1	基于主题图的旅游文献组织方法研究	(175)
11.1.1	旅游文献对象分析	(175)
11.1.2	信息环境与信息需求分析	(177)
11.1.3	旅游文献主题图基本元素分析	(178)
11.2	基于主题图的旅游文献知识管理模型研究	(188)
11.2.1	领域结构分析与模型	(189)
11.2.2	TM-TDKM 基本定义	(190)
11.2.3	TDKM 知识模型定义	(191)
11.2.4	TDKS 知识服务定义	(192)
11.3	基于主题图的信息检索模型研究	(194)
11.3.1	模型基本定义	(195)
11.3.2	模型实现	(196)
11.3.3	验证分析及比较	(200)
	参考文献	(201)
第 12 章	结论	(204)
12.1	研究结果	(204)
12.1.1	相关研究结论	(204)
12.1.2	研究创新之处	(205)
12.1.3	研究不足之处	(206)
12.2	未来研究方向	(207)
	后 记	(208)

主题图基础理论篇

第1章 绪论

1.1 主题图发展历程

1.1.1 主题图引出

有人说：“一本没有索引的书就像一个国家没有地图一样。”我们最直接的经验就是：如果要驾车从A地到B地，若是没有地图，我们将会费尽周折。毫无疑问，任何形式的地图，将是我们能够顺利到达目的地最不可或缺的东西。

同理，我们要在书中寻找一条信息，除非你喜欢从头到尾地享受阅读给你带来的快乐，否则一个好的索引是你必不可少的助手。对于在一本书中寻找一条特定的信息而言，一个好的索引是非常有价值的。索引可以看做是在对书中的内容作了仔细研究后专门制作的地图。一般情况下，索引除了可以完整地列出所有出现在该书中的重要关键词之外，更重要的是它也可以指出关键词之间的联系，即可以建立关键词之间“见(See)”以及“参见(See Also)”之参照关系，将具有关联性的关键词加以联结，读者就能够按图索骥，找到真正需要的信息。正如拉里·布鲁诺(Larry Bonura)所指出那样：“索引就是将文献的主题做成图表，从而呈现给读者一个简练、准确的地图。”

在信息时代，直到现在还没有创造出一种类似于索引的东西，来帮助用户从大量的数字信息中查找出自己所需要的特定信息。的确，现在人们已经学会在文档处理中标记关键词，并利用它来自动生成索引。然而这种形式的索引还只能局限在单一文档模式之中。对于数字化的信息空间则完全不同。例如，我们大家都非常熟悉的互联网(World Wide Web)中，由于文档个体形式之间的差异消失了，并且对索引的要求也不仅局限在文档之内，而是要能跨越不同的文档。因此，反过来要求不仅要具备索引合并功能，还应能创建基于用户视图的信息资源(User-Defined Views of Information)^[1]，这样传统的索引技术就会受到挑战。

由于纸质文献的索引结构和内容是固定不变的，用户关注的也仅仅是单个文献内部的特定信息。而对于数字信息，用户希望能够按照自己的需要检索出多篇文献，在某些情况下甚至还要包括大量的信息集合。这就要求我们能够按照用户定义的信息特征创建和合并索引，以适应检索的需要。这样，传统的索引技术就无法满足用户的这种需求变化。

其实，在文献管理领域，早在几十年前就遇到了上述问题，但是解决这个问题的方